

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension .rup) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2005, 14:39:13 ; Search time 169 Seconds
(without alignments)
1375.402 Million cell updates/sec

Title: US-10-712-479-2

Perfect score: 3167

Sequence: 1 MASMSMTWTGALITPCAA.....LLSVGVGIYLLPNRRHHHH 601

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3167	100.0	601	ADO19021	ADO19021 Full-leng
2	3034.5	95.8	580	ADO19023	ADO19023 C-termina
3	3027	95.6	3010	AAR30616	AAR30616 Polypepti
4	3016	95.2	591	AAR01679	AAR01679 NS5B prot
5	3016	95.2	2201	AAR01680	AAR01680 HCV NS2-N
6	3016	95.2	3010	AAR20111	AAR20111 Non-A, no
7	3016	95.2	3010	AAR20091	AAR20091 Non-A, no
8	3016	95.2	3010	ADF88597	ADF88597 Hepatitis
9	3015	95.2	3010	AAW98022	AAW98022 Infectiou
10	3015	95.2	3010	AAW31170	AAW31170 Amino aci
11	3015	95.2	3010	ADO36227	ADO36227 Hepatitis
12	3015	95.2	3010	ADO79401	ADO79401 Hepatitis
13	3015	95.2	592	ADO36215	ADO36215 Hepatitis
14	3014	95.2	592	ADO79389	ADO79389 Hepatitis
15	3014	95.2	3010	AAY06423	AAY06423 Non-A, no
16	3013	95.1	3010	AAR68864	AAR68864 Hepatitis
17	3012	95.1	591	AAW60130	AAW60130 Hepatitis
18	3012	95.1	591	AAO26416	AAO26416 Protein d
19	3012	95.1	1985	ABU09575	ABU09575 HCV Met-N
20	3012	95.1	3010	AAR68622	AAR68622 HCV prote
21	3011	95.1	591	AAG79560	AAG79560 HCV NS5B
22	3011	95.1	621	AAG79556	AAG79556 HCV decre
23	3011	95.1	621	AAE18903	AAE18903 Hepatitis
24	3011	95.1	621	ABP71699	ABP71699 HCV NS5B
25	3011	95.1	621	ABR41892	ABR41892 Hepatitis

26	3011	95.1	621	6	ABP97967	Abp97967 Amino aci
27	3011	95.1	627	5	AAG79561	Aag79561 HCV decre
28	3011	95.1	627	5	ADG19078	Adg19078 HCV NSSB
29	3011	95.1	3010	2	AAR82694	Aar82694 Partial H
30	3005	94.9	1985	5	AAO18001	Aao18001 Hepatitis
31	3005	94.9	1985	5	AAE15729	Aae15729 Hepatitis
32	3005	94.9	1985	5	AAE15731	Aae15731 Hepatitis
33	3005	94.9	1985	5	AAE15720	Aae15720 Hepatitis
34	3005	94.9	1985	5	AAE15717	Aae15717 Hepatitis
35	3005	94.9	1985	5	AAE15727	Aae15727 Hepatitis
36	3005	94.9	1985	5	AAE15728	Aae15728 Hepatitis
37	3005	94.9	1985	5	AAE15722	Aae15722 Hepatitis
38	3005	94.9	1985	5	AAE15730	Aae15730 Hepatitis
39	3005	94.9	1985	8	ADJ57846	Adj57846 HCV repli
40	3005	94.9	2201	5	ABG30591	Abg30591 Hepatitis
41	3005	94.9	2201	5	ABG30600	Abg30600 Hepatitis
42	3005	94.9	2201	5	ABG30581	Abg30581 Hepatitis
43	3005	94.9	2201	5	ABG30586	Abg30586 Hepatitis
44	3005	94.9	2201	5	ABG30593	Abg30593 Hepatitis
45	3005	94.9	2201	5	ABG30582	Abg30582 Hepatitis

ALIGNMENTS

RESULT 1
ADO19021
ID ADO19021 standard; protein; 601 AA.
XX ADO19021;
XX
XX
DT 12-AUG-2004 (first entry)
XX
DE Full-length HCV NS5B polymerase, FL NS5B.
XX

RNA polymerase activity; continuous-read assay;
Hepatitis C virus infection; HCV infection; polymerase; NS5B;
bone mineral disease; osteoporosis; carcinoma; cardiovascular disease;
diabetes; ocular disorder; renal dysfunction; lymphoma;
lymphoproliferative disorder; metabolic disorder; arthritis;
sleep disorder; thyroid disorder; antiinflammatory; hepatotropic;
viral; osteopathic; cytostatic; antidiabetic; ophthalmological;
nephrotropic; antiarthritic; enzyme.

Hepatitis C virus.

WO200404228-A2.

27-MAY-2004.

13-NOV-2003; 2003WO-US036465.

13-NOV-2002; 2002US-0425981P.

(PFIZ) PFIZER INC.

Yagi Y, Sheets MP, Wells PA, Shelly JA, Poorman RA, Epps DB;
Morgan AG;

WPI; 2004-420337/39.

N-PSDB; ADO19020.

Detecting RNA polymerase activity in a continuous-read manner, useful in
treating osteoporosis, carcinomas, cardiovascular diseases, ocular
disorders or arthritis, by contacting an RNA polymerase with an
oligonucleotide template.

Claim 5; SEQ ID NO 2; 46pp; English.

The present invention relates to a method for detecting RNA polymerase
activity in a continuous-read manner. The method comprises contacting an
RNA polymerase with an oligonucleotide template in a reaction mixture
comprising an assay buffer, under conditions in which the RNA polymerase

CC is active, adding a fluorescent dye capable of binding double-stranded
 CC nucleic acid molecules to the reaction mixture. The RNA polymerase is the Hepatitis
 CC fluorescence of the reaction mixture. The RNA polymerase is the Hepatitis
 CC C virus (HCV) polymerase, NS5B. The method is useful for treating HCV
 CC infection, bone mineral diseases like osteoporosis, carcinomas,
 CC cardiovascular diseases, diabetes, ocular disorders, renal dysfunction,
 CC lymphomas, lymphoproliferative disorders, metabolic disorders, arthritis,
 CC sleep disorders and thyroid disorders. The present sequence represents
 CC full-length HCV NS5B polymerase (designated FL NS5B).

XX Sequence 601 AA;
 SQ Query Match 100.0%; Score 3167; DB 8; Length 601;
 Best Local Similarity 100.0%; Pred. No. 4e-294;
 Matches 601; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MASMSTYTTGALITPCAAEESKLPINALSNSLLRHHNLVYSTRSASLRQKVTDFR 60
 DB 1 MASMSTYTTGALITPCAAEESKLPINALSNSLLRHHNLVYSTRSASLRQKVTDFR 60
 QY 61 LOVLDDHYRDLKEMKAKASTVKAALLSVEAECKLTPPHSAKSKFGYGAQDVRLSSRAV 120
 DB 61 LOVLDDHYRDLKEMKAKASTVKAALLSVEAECKLTPPHSAKSKFGYGAQDVRLSSRAV 120
 QY 121 NHIRSVKDLLEDTDTPITQTTIMAKNEVFCVQPEKGRKPARLIVFPDGLGVRCCKMAY 180
 DB 121 NHIRSVKDLLEDTDTPITQTTIMAKNEVFCVQPEKGRKPARLIVFPDGLGVRCCKMAY 180
 QY 181 DVVSTLPOAVNGSSYGFQYSPKQRFVLTWTKAKKCPMGFSYDTRCFDSTVTENDIRVE 240
 DB 181 DVVSTLPOAVNGSSYGFQYSPKQRFVLTWTKAKKCPMGFSYDTRCFDSTVTENDIRVE 240
 QY 241 ESIYQCCDLAPPEARQAIRSLRTERLYVGGPMTNSKGQNGYRRCRASCGLVLTSSCGNTLTCTY 300
 DB 241 ESIYQCCDLAPPEARQAIRSLRTERLYVGGPMTNSKGQNGYRRCRASCGLVLTSSCGNTLTCTY 300
 QY 301 LKAAACRAAKLQDCTMLVNGDDLVIWICESAGTQEDAAASLRVFTTEAMTRYAPPDPPQP 360
 DB 301 LKAAACRAAKLQDCTMLVNGDDLVIWICESAGTQEDAAASLRVFTTEAMTRYAPPDPPQP 360
 QY 361 EYDLELTSSNSVSAHDASGKRVYLLTRDPTVPLARAAMETARHTPVNSWLGNIIMYA 420
 DB 361 EYDLELTSSNSVSAHDASGKRVYLLTRDPTVPLARAAMETARHTPVNSWLGNIIMYA 420
 QY 421 PTLWARMILMTHFFSILLAQBLEKALDCQYAGACYSIEPLDLQITLERHGLSAFSLHS 480
 DB 421 PTLWARMILMTHFFSILLAQBLEKALDCQYAGACYSIEPLDLQITLERHGLSAFSLHS 480
 QY 481 YSPGEINRVASCLRLKGLVPPRLVRHRRARSVRKLLSQGGRALCGKYLFWAVRTKLL 540
 DB 481 YSPGEINRVASCLRLKGLVPPRLVRHRRARSVRKLLSQGGRALCGKYLFWAVRTKLL 540
 QY 541 TPIPAASRLDLSGFWAGYSGGDIYHLSRARPRFMLCLLLSVGVGIYLLPNRHHHH 600
 DB 541 TPIPAASRLDLSGFWAGYSGGDIYHLSRARPRFMLCLLLSVGVGIYLLPNRHHHH 600
 QY 601 H 601
 DB 601 H 601

RESULT 2
 ADO19023
 ID ADO19023 standard; protein; 580 AA.
 XX AC ADO19023;
 XX AC ADO19023;
 DT 12-AUG-2004 (first entry)
 XX C-terminally truncated HCV NS5B polymerase, Cdelta21 NS5B.
 DE RNA polymerase activity; continuous-read assay;
 KW Hepatitis C virus infection; HCV infection; polymerase; NS5B;

KW bone mineral disease; osteoporosis; carcinoma; cardiovascular disease;
 KW diabetes; ocular disorder; renal dysfunction; lymphoma;
 KW lymphoproliferative disorder; metabolic disorder; arthritis;
 KW sleep disorder; thyroid disorder; antiinflammatory; hepatotropic;
 KW virucide; osteopathic; cytostatic; antidiabetic; ophthalmological;
 KW nephrotropic; antiarthritic; enzyme.
 OS Hepatitis C virus.
 XX WO2004044228-A2.
 FN 27-MAY-2004.
 XX 13-NOV-2003; 2003WO-US036465.
 XX 13-NOV-2002; 2002US-0425981P.
 XX (PFIZ) PFIZER INC.
 XX Yagi Y, Sheets MP, Wells PA, Shelly JA, Poorman RA, Epps DE;
 FI Morgan AG;
 XX WPI; 2004-420337/39.
 DR N-PSDB; ADO19022.
 XX Detecting RNA polymerase activity in a continuous-read manner, useful in
 XX treating osteoporosis, carcinomas, cardiovascular diseases, ocular
 PT disorders or arthritis, by contacting an RNA polymerase with an
 PT oligonucleotide template.
 XX Claim 7; SEQ ID NO 4; 46pp; English.

XX The present invention relates to a method for detecting RNA polymerase
 CC activity in a continuous-read manner. The method comprises contacting an
 CC RNA polymerase with an oligonucleotide template in a reaction mixture
 CC comprising an assay buffer, under conditions in which the RNA polymerase
 CC is active, adding a fluorescent dye capable of binding double-stranded
 CC nucleic acid molecules to the reaction mixture, and measuring the
 CC fluorescence of the reaction mixture. The RNA polymerase is the Hepatitis
 CC C virus (HCV) polymerase, NS5B. The method is useful for treating HCV
 CC infection, bone mineral diseases like osteoporosis, carcinomas,
 CC cardiovascular diseases, diabetes, ocular disorders, renal dysfunction,
 CC lymphomas, lymphoproliferative disorders, metabolic disorders, arthritis,
 CC sleep disorders and thyroid disorders. The present sequence represents C-
 CC terminally truncated HCV NS5B polymerase (designated Cdelta21 NS5B).

XX Sequence 580 AA;

Query Match 95.8%; Score 3034.5; DB 8; Length 580;
 Best Local Similarity 96.5%; Pred. No. 1.9e-281;
 Matches 580; Conservative 0; Mismatches 0; Indels 21; Gaps 1;

QY 1 MASMSTYTTGALITPCAAEESKLPINALSNSLLRHHNLVYSTRSASLRQKVTDFR 60
 DB 1 MASMSTYTTGALITPCAAEESKLPINALSNSLLRHHNLVYSTRSASLRQKVTDFR 60
 QY 61 LOVLDDHYRDLKEMKAKASTVKAALLSVEAECKLTPPHSAKSKFGYGAQDVRLSSRAV 120
 DB 61 LOVLDDHYRDLKEMKAKASTVKAALLSVEAECKLTPPHSAKSKFGYGAQDVRLSSRAV 120
 QY 121 NHIRSVKDLLEDTDTPITQTTIMAKNEVFCVQPEKGRKPARLIVFPDGLGVRCCKMAY 180
 DB 121 NHIRSVKDLLEDTDTPITQTTIMAKNEVFCVQPEKGRKPARLIVFPDGLGVRCCKMAY 180
 QY 181 DVVSTLPOAVNGSSYGFQYSPKQRFVLTWTKAKKCPMGFSYDTRCFDSTVTENDIRVE 240
 DB 181 DVVSTLPOAVNGSSYGFQYSPKQRFVLTWTKAKKCPMGFSYDTRCFDSTVTENDIRVE 240
 QY 241 ESIYQCCDLAPPEARQAIRSLRTERLYVGGPMTNSKGQNGYRRCRASCGLVLTSSCGNTLTCTY 300
 DB 241 ESIYQCCDLAPPEARQAIRSLRTERLYVGGPMTNSKGQNGYRRCRASCGLVLTSSCGNTLTCTY 300
 QY 301 LKAAACRAAKLQDCTMLVNGDDLVIWICESAGTQEDAAASLRVFTTEAMTRYAPPDPPQP 360

Db 301 LKAAACRAAKLQDCTMLVNGDDLVVICESAGTQBDAAASLRVFTAMTRYSPAPGDPQP 360
QY 361 EYDLELITSCSSNVSAHDASGKRVYLLTRDPTVPLARAAMETARHTPVNSWLGNIIMYA 420
Db 361 EYDLELITSCSSNVSAHDASGKRVYLLTRDPTVPLARAAMETARHTPVNSWLGNIIMYA 420
QY 421 PTLWARMILMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIRLHGLSAPLSLHS 480
Db 421 PTLWARMILMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIRLHGLSAPLSLHS 480
QY 481 YSPGGINRVASCLRLKGLVPPPLRVWHRARSVRAKLLSQGGRAAICGKYLFWAVRTKL 540
Db 481 YSPGGINRVASCLRLKGLVPPPLRVWHRARSVRAKLLSQGGRAAICGKYLFWAVRTKL 540
QY 541 TPIPAASRLDLSGWFVAGYSGGDIYHSLSRAPRPFWMFLCLLLSVGVGVIYLLPNRHHHH 600
Db 541 TPIPAASRLDLSGWFVAGYSGGDIYHSLSRAPRPFWMFLCLLLSVGVGVIYLLPNRHHHH 600
QY 601 H 601
Db 580 H 580

RESULT 3
AAAR30616
ID AAR30616 standard; protein; 3010 AA.

AAAR30616;

27-AUG-2003 (revised)
25-MAR-2003 (revised)
19-MAY-1993 (first entry)

Polypeptide coded by Korean HCV full cDNA sequence LBCL.

KHCV-LBCL; diagnosis; vaccine.

Hepatitis C virus.

EP521318-A2.

07-JAN-1993.

10-JUN-1992; 92EP-00109753.

10-JUN-1991; 91KR-00009510.

06-AUG-1991; 91KR-00013601.

(LUCK-) LUCKY LTD.

Cho JM, Lee YB, Park YW, Lim KJ, Choi DY, So HS, Kim CH;

Kim ST, Yang JY;

WPI; 1993-001883/01.

N-PSDB; AAQ33282.

DNA and polypeptide(s) from a new type of hepatitis C virus (KHCV) - for

diagnosing and vaccinating against KHCV infections.

Disclosure; Fig 2; 119pp; English.

The polypeptide is that encoded by the full cDNA sequence of Korean hepatitis C virus (KHCV) cDNA, KHCV-LBCL. It or its fragments may be used in a specific and accurate method for detecting KHCV antibodies in the serum of hepatitis C patients. Antibodies directed against these polypeptides are useful for the purification of KHCV antigens and for the development of an improved diagnostic to detect KHCV antigens in a sample. The polypeptides may also be used in a vaccine for treatment and prevention of KHCV infection at a dosage of 5-200 ug/peptide. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS field.)

SQ Sequence 3010 AA;
Query Match 95.6%; Score 3027; DB 2; Length 3010;
Best Local Similarity 96.8%; Pred. No. 1.3e-279;
Matches 572; Conservative 9; Mismatches 10; Indels 0; Gaps 0;
QY 5 SMSYTWTCALITPCAAEESKLPINALSNLRLHHNLVYSTSRSSASLRQKVKVTFORLQVL 64
Db 2420 SMSYTWTCALITPCAAEESKLPINPLSNLRLHHNMVATTSSAGLRQKVKVTFORLQVL 2479
QY 65 DDHYRDVLKEMKAKASTVKALLSVEEACKLTPPHSAKSKFGYGADEVSRSLSSRAVNHIR 124
Db 2480 DDHYRDVLKEMKAKASTVKALLSVEEACKLTPPHSAKSKFGYGADEVSRSLSSRAVNHIR 2539
QY 125 SVWKDLLEDTTPIOTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 184
Db 2540 SVWKDLLEDTTPIOTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 2599
QY 185 TLPOAVMGSSYGFQYSPKQVFEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
Db 2600 TLPOAVMGSSYGFQYSPKQVFEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 2659
QY 245 QCCDLAPEARQAIRSLTERLYVGGPMVNSKGNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
Db 2660 QCCDLAPEARQAIRSLTERLYVGGPMVNSKGNCGYRRCRASGVLTTSCGNTLTCYLKAT 2719
QY 305 AACRAAKLQDCTMLVNGDDLVVICESAGTQBDAAASLRVFTAMTRYSPAPGDPPEYDL 364
Db 2720 AACRAAKLQDCTMLVNGDDLVVICESAGTQBDAAASLRVFTAMTRYSPAPGDPPEYDL 2779
QY 365 ELITSCSSNVSAHDASGKRVYLLTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTLM 424
Db 2780 ELITSCSSNVSAHDASGKRVYLLTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTLM 2839
QY 425 ARMLMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIRLHGLSAPLSLHSYSPG 484
Db 2840 ARMLMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIRLHGLSAPLSLHSYSPG 2899
QY 485 EINRVASCLRLKGLVPPPLRVWHRARSVRAKLLSQGGRAAICGKYLFWAVRTKLTPIP 544
Db 2900 EINRVASCLRLKGLVPPPLRVWHRARSVRAKLLSQGGRAATCGKYLFWAVRTKLTPIP 2959
QY 545 AASRLDLSGWFVAGYSGGDIYHSLSRAPRPFWMFLCLLLSVGVGVIYLLPNR 595
Db 2960 AASRLDLSGWFVAGYSGGDIYHSLSRAPRPFWMFLCLLLSVGVGVIYLLPNR 3010
RESULT 4
AAW01679
ID AAW01679 standard; protein; 591 AA.
XX AC AAW01679;
XX DT 17-OCT-2003 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 19-AUG-1997 (first entry)
XX DE NSB protein (residues 2420-3010 of the HCV polyprotein).
XX KW Hepatitis C virus; HCV; non-structural protein; NSB; polyprotein;
XX KW RNA-dependent RNA polymerase; RdRp; terminal nucleotidyl transferase;
XX KW Tntase; method; assay; in vitro activity; therapy; inhibitor.
XX OS Hepatitis C virus; strain BK.
XX FN WO9637619-A1.
XX PD 28-NOV-1996.
XX PF 24-MAY-1996; 96WO-IT000106.
XX PR 25-MAY-1995; 95IT-RM000343.
XX XX

(RICE-) 1ST RICERCHE BIOL MOLECOLARE ANGELETTI.

De Francesco R, Tomei L, Behrens S;

WPI; 1997-021225/02.

Reproducing enzymatic activities of HCV in vitro - using sequences contg. NS5B for RNA-dependent RNA polymerase and terminal nucleotidyl transferase activities.

Claim 1; Page 24-26; 49pp; English.

A novel method for reproducing in vitro the RNA-dependent RNA polymerase (RdRp) activity or the terminal nucleotidyl transferase (TNTase) activity encoded by hepatitis C virus (HCV), is characterised in that sequences contg. NS5B are used in the mixture. The method is used for assaying in vitro the activities of RdRp and TNTase encoded by HCV in order to identify, for therapeutic purposes, compounds that inhibit these enzymatic activities and therefore might interfere with the replication of the HCV. The present sequence comprises amino acids 2420-3010 of the HCV polyprotein and corresponds to the NS5B protein. cDNA encoding this protein was cloned between the BamHI and HindIII sites of pBlueBacIII to form pBac58. Another expression plasmid, pBac25 (containing cDNA encoding amino acids 810-3010 of HCV; see AAW01680) was also constructed. Extracts of Bac25- or Bac58-infected Sf9 cells contain a novel magnesium-dependent enzymatic activity that catalyses de novo RNA synthesis. This activity was shown to be dependent on the presence of RNA, but independent of an added primer or of the origin of the input RNA molecule. As the products generated by extracts of Sf9 cells infected with either Bac25 or Bac58 appeared to be identical, the experiments indicated that the observed RdRp activity is encoded by the HCV NS5B protein. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 17-OCT-2003 to standardise OS field)

Sequence 591 AA;

Query Match 95.2%; Score 3016; DB 2; Length 591;
Best Local Similarity 96.6%; Pred. No. 1.2e-279;
Matches 571; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

QY 5 SMSYTWGALITPCAARESKLPINALNSLRHNHNVYTSRSASLRQKVKVTFDRQLVL 64
DB 1 SMSYTWGALITPCAARESKLPINALNSLRHNHNVYATTSRSAGLRQKVKVTFDRQLVL 60

QY 65 DDHVRDLKEMKAKSTVKALLSVEERACKLTPPHSAKSKFGYGAQVRNLSKAVNHIR 124
DB 61 DDHVRDLKEMKAKSTVKALLSVEERACKLTPPHSAKSKFGYGAQVRNLSKAVNHII 120

QY 125 SVWKDLLEDTPITQTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMAKYDVS 184
DB 121 SVWKDLLEDTPITQTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMAKYDVS 180

QY 185 TLPQAVMGSSYGFQYSPQQRVEFLVNTWKAKKPMGFSYDFRCFSTVTENDIRVESIY 244
DB 181 TLPQAVMGSSYGFQYSPQQRVEFLVNTWKSKNPMGFSYDFRCFSTVTENDIRVESIY 240

QY 245 QCCDLAPARQAIKSLTERLVGGPMTNKQKNGCYRCRAGSVLTSCGNLTFCYLKAA 304
DB 241 QCCDLAPARQAIKSLTERLVGGPMTNKQKNGCYRCRAGSVLTSCGNLTFCYLKAS 300

QY 305 AACRAAKLQDCTMLVNGDVLVVCESACTQEDAAASLRVFTAMTRYSAAPPDPPQPEYDL 364
DB 301 AACRAAKLQDCTMLVNGDVLVVCESACTQEDAAASLRVFTAMTRYSAAPPDPPQPEYDL 360

QY 365 ELITSCSNNSVAHDAGKRVYVTRDPTVPLARAAMETARHTPVNSWLGNIIMVAPTLW 424
DB 361 ELITSCSNNSVAHDAGKRVYVTRDPTVPLARAAMETARHTPVNSWLGNIIMVAPTLW 420

QY 425 ARMLIMTHFFSILLAQOEKALDQCIYGCACVSTIEPLDLPQIERLHCLSAFSLHSPG 484
DB 421 ARMLIMTHFFSILLAQOEKALDQCIYGCACVSTIEPLDLPQIERLHCLSAFSLHSPG 480

QY 485 EINRVASCLRLKGVPPPLRVWRHRRARSVRALLSQGGRAATCGKYLEFNWAVTKLKLTPIP 544
DB 481 EINRVASCLRLKGVPPPLRVWRHRRARSVRALLSQGGRAATCGKYLEFNWAVTKLKLTPIP 540

DB 481 EINRVASCLRLKGVPPPLRVWRHRRARSVRALLSQGGRAATCGKYLEFNWAVTKLKLTPIP 540
QY 545 AASRLDLSGWFVAGYSGGDIYHLSLRARPRWFMCLCLLLSVGVGIYLLPNR 595
DB 541 AASRLDLSGWFVAGYSGGDIYHLSLRARPRWFMCLCLLLSVGVGIYLLPNR 591

RESULT 5

AAW01680
ID AAW01680 standard; protein; 2201 AA.

XX AC AAW01680;

XX DT 17-OCT-2003 (revised)

XX DT 25-MAR-2003 (revised)

XX DT 19-AUG-1997 (first entry)

XX DE HCV NS2-NS5B non-structural protein.

XX KW Hepatitis C virus; HCV; non-structural protein; NS5B; polyprotein;

XX KW RNA-dependent RNA polymerase; RdRp; terminal nucleotidyl transferase;

XX KW TNTase; method; assay; in vitro activity; therapy; inhibitor.

XX OS Hepatitis C virus; strain BK.

XX PN W09637619-A1.

XX PD 28-NOV-1996.

XX PF 24-MAY-1996; 96WO-IT000106.

XX PR 25-MAY-1995; 95IT-RM000343.

XX PA (RICE-) 1ST RICERCHE BIOL MOLECOLARE ANGELETTI.

XX PI De Francesco R, Tomei L, Behrens S;

XX DR WPI; 1997-021225/02.

PT Reproducing enzymatic activities of HCV in vitro - using sequences contg. NS5B for RNA-dependent RNA polymerase and terminal nucleotidyl transferase activities.

XX Example 1; Page 26-34; 49pp; English.

CC A novel method for reproducing in vitro the RNA-dependent RNA polymerase (RdRp) activity or the terminal nucleotidyl transferase (TNTase) activity encoded by hepatitis C virus (HCV), is characterised in that sequences contg. NS5B are used in the mixture. The method is used for assaying in vitro the activities of RdRp and TNTase encoded by HCV in order to identify, for therapeutic purposes, compounds that inhibit these enzymatic activities and therefore might interfere with the replication of the HCV. The present sequence comprises amino acids 810-3010 of the HCV polyprotein and corresponds to NS2-NS5B proteins. cDNA encoding this protein was cloned between the NcoI and HindIII sites of pBlueBacIII to form pBac25. Another expression plasmid, pBac58 (containing cDNA encoding amino acids 2420-3010 of HCV; see AAW01679) was also constructed. Extracts of Bac25- or Bac58-infected Sf9 cells contain a novel magnesium-dependent enzymatic activity that catalyses de novo RNA synthesis. This activity was shown to be dependent on the presence of RNA, but independent of an added primer or of the origin of the input RNA molecule. As the products generated by extracts of Sf9 cells infected with either Bac25 or Bac58 appeared to be identical, the experiments indicated that the observed RdRp activity is encoded by the HCV NS5B protein. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 17-OCT-2003 to standardise OS field)

XX SQ Sequence 2201 AA;

Query Match 95.2%; Score 3016; DB 2; Length 2201;

Best Local Similarity 96.6%; Pred. No. 9.1e-279;

Matches 571; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

QY 5 SMSYTWTCALITPCAARESKLPINALNSLLRHHNLVYSTTSRSASLRQKKVTFDRLOVL 64
 Db 1611 SMSYTWTCALITPCAARESKLPINALNSLLRHHNLVYATTSRSAGLRQKKVTFDRLOVL 1670
 QY 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAQVRSLSRAVNHIR 124
 Db 1671 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAQVRSLSRAVNHIR 1730
 QY 125 SVWKDLLEDTPITQTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMAIYDVVS 184
 Db 1731 SVWKDLLEDTPITQTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMAIYDVVS 1790
 QY 185 TLPOAVMGSSYGFQYSPKQRFVFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
 Db 1791 TLPOAVMGSSYGFQYSPKQRFVFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 1850
 QY 245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQCCYRRCRASGVLTTCGNTLTCLYKAA 304
 Db 1851 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQCCYRRCRASGVLTTCGNTLTCLYKAA 1910
 QY 305 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAASLRVFTTEAMTRYSPAPGDPPOPEYDL 364
 Db 1911 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAASLRVFTTEAMTRYSPAPGDPPOPEYDL 1970
 QY 365 ELITSCSSNVSAHDASGRVYLLTRDPTVPLARAAMETARHTPVSMLGNIIMYAPTLM 424
 Db 1971 ELITSCSSNVSAHDASGRVYLLTRDPTVPLARAAMETARHTPVSMLGNIIMYAPTLM 2030
 QY 425 ARMILMTHFFSILLAOEQLKALDCQIYGCACYSIEPLDLPOIIRLHGLSAPLSHSYSPG 484
 Db 2031 ARMILMTHFFSILLAOEQLKALDCQIYGCACYSIEPLDLPOIIRLHGLSAPLSHSYSPG 2090
 QY 485 EINRVASCLRLKGLVPLRWRHRRARSVRKALLSQGGRAAICGKYLFWNAVTKLKTPIIP 544
 Db 2091 EINRVASCLRLKGLVPLRWRHRRARSVRKALLSQGGRAATCGKYLFWNAVTKLKTPIIP 2150
 QY 545 AASRLDLGWFVAGYSGGDIYHLSLRARPRWFMCLCLLLSVGVGYLLPNR 595
 Db 2151 AASRLDLGWFVAGYSGGDIYHLSLRARPRWFMCLCLLLSVGVGYLLPNR 2201

RESULT 6
 AAR20111
 ID AAR20111 standard; protein; 3010 AA.
 XX AC AAR20111;
 XX AC AAR20111;
 DT 25-MAR-2003 (revised)
 DT 01-MAY-1992 (first entry)
 XX DE Non-A, non-B viral genome product.
 XX KW NANBV; vaccine; immunodiagnosis; antigen; antibody.
 XX OS Non-A.
 XX OS Non-B hepatitis virus.
 FH Key Location/Qualifiers
 FT Protein 1..115
 FT /label= C
 FT /note= "core protein"
 FT 116..191
 FT /label= M
 FT /note= "matrix protein"
 FT 192..389
 FT /label= E
 FT /note= "envelope protein"
 FT 390..729
 FT /label= NS1
 FT 730..1006
 FT /label= NS2
 FT 1007..1614
 FT /label= NS3

FT Protein 1615..1862
 FT /label= NS4a
 FT 1863..2012
 FT /label= NS4b
 FT 2013..3010
 FT /label= NS5
 XX EP463848-A.
 XX 02-JAN-1992.
 XX 25-JUN-1991; 91EP-00305717.
 XX 25-JUN-1990; 90JP-00167466.
 PR 31-AUG-1990; 90JP-00230921.
 PR 09-NOV-1990; 90JP-00305605.
 PR 28-DEC-1990; 90US-00635451.
 PR 08-MAY-1991; 91JP-00132090.
 PR 14-MAY-1991; 91JP-00138493.
 XX (OSAU) UNIV OSAKA.
 PA (REMI-) RES FOUND MICROBIAL.
 XX WPI; 1992-009412/02.
 DR N-PSDB; AAQ20268.
 XX Non-A, non-B hepatitis virus (NANBV) particles - as vaccines, immuno-
 PT diagnostics and screening agents for NANBV, and to remove NANBV from
 FT blood.
 XX Disclosure; Fig 2; 89pp; English.
 XX The sequence (SEQ ID NO 1) was deduced from several overlapping from a
 CC library prep'd. from NANBV RNA. Antigenic polypeptides from the sequence
 CC can be used as immunoscreening reagents, for screening donated blood, and as
 CC immunogens for vaccine prodn. Antibodies raised to the peptides can be
 CC used in immunoassays to detect or quantify NANBV antigens in liver tissue
 CC and blood. Preferred poly- peptides include residues 1-30, -115, or 2012;
 CC 47-77; 116-191; 192-207 or -298; 230-238 or -263; 287-300; 293-330; 390-
 CC 729; 730-1005; 1006-1614; 1384-1414; 1615-1862; 1737-1767; 1863-2012;
 CC and 2013-3010. The sequence is also disclosed in EP-464287 (SEQ ID NO 1).
 CC See AAR20091 for details of this specification. (Updated on 25-MAR-2003
 CC to correct PD field.) (Updated on 25-MAR-2003 to correct PA field.)
 XX SQ Sequence 3010 AA;
 Query Match 95.2%; Score 3016; DB 2; Length 3010;
 Best Local Similarity 96.6%; Pred. No. 1.5e-278;
 Matches 571; Conservative 11; Mismatches 9; Indels 0; Gaps 0;
 QY 5 SMSYTWTCALITPCAARESKLPINALNSLLRHHNLVYSTTSRSASLRQKKVTFDRLOVL 64
 Db 2420 SMSYTWTCALITPCAARESKLPINALNSLLRHHNLVYATTSRSAGLRQKKVTFDRLOVL 2479
 QY 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAQVRSLSRAVNHIR 124
 Db 2480 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAQVRSLSRAVNHIR 2539
 QY 125 SVWKDLLEDTPITQTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMAIYDVVS 184
 Db 2540 SVWKDLLEDTPITQTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMAIYDVVS 2599
 QY 185 TLPOAVMGSSYGFQYSPKQRFVFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
 Db 2600 TLPOAVMGSSYGFQYSPKQRFVFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 2659
 QY 245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQCCYRRCRASGVLTTCGNTLTCLYKAA 304
 Db 2660 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQCCYRRCRASGVLTTCGNTLTCLYKAA 2719
 QY 305 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAASLRVFTTEAMTRYSPAPGDPPOPEYDL 364
 Db 2720 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAASLRVFTTEAMTRYSPAPGDPPOPEYDL 2779

QY	365	ELITSCSSNVSV	AHDA	SGKRV	YVYL	TRDPT	VP	LARA	AWET	ARHT	TPV	PN	SWL	GNII	IM	YAP	TLW	424										
Db	2780	ELITSCSSNVSV	AHDA	SGKRV	YVYL	TRDPT	TP	LARA	AWET	ARHT	TPV	PN	SWL	GNII	IM	YAP	TLW	2839										
QY	425	ARMILMTHFFS	II	LAQ	EQLE	KALD	QC	YI	GAC	YSIE	PLD	L	PQ	II	IER	L	HGL	SFA	SLHS	YSPG	484							
Db	2840	ARMILMTHFFS	II	LAQ	EQLE	KALD	QC	YI	GAC	YSIE	PLD	L	PQ	II	IER	L	HGL	SFA	SLHS	YSPG	2899							
QY	485	EINRVASCL	RKL	GV	PP	LV	R	VR	H	R	ARS	VR	AK	L	SGG	GRA	A	CG	KYL	FN	WA	V	RT	KL	K	LT	PP	544
Db	2900	EINRVASCL	RKL	GV	PP	LV	R	VR	H	R	ARS	VR	AK	L	SGG	GRA	A	CG	KYL	FN	WA	V	RT	KL	K	LT	PP	2959
QY	545	AASRLDLS	GH	FW	VAG	YSG	GD	I	YH	SL	SR	AP	R	PN	FW	M	CL	L	L	L	S	V	G	YI	Y	L	PN	595
Db	2960	AASRLDLS	GH	FW	VAG	YSG	GD	I	YH	SL	SR	AP	R	PN	FW	M	CL	L	L	S	V	G	YI	Y	L	PN	3010	

RESULT 7

AAR20091
 ID AAR20091 standard; protein; 3010 AA.

XX Hepatitis C virus NS3 gene protein, SEQ ID No 6.
XX liver cancer; hepatitis-C virus; NS3 gene; carcinogenicity.
XX Hepatitis C virus.
XX JP2003210181-A.
XX 29-JUL-2003.
XX 30-MAY-2002; 2002JP-00158335.
XX 16-NOV-2001; 2001JP-00352443.
XX (SHIN-) ZH SHINSANGYO SOZO KENKYU KIKO.
XX WPI: 2003-819836/77.
XX N-PSDB; ADF88596.

DR Diagnosing liver cancer, involves amplifying amino terminal region of
PT hepatitis-C virus gene using predetermined primer and determining
PT hepatitis-C virus in base sequence of obtained DNA fragment.
XX Disclosure; SEQ ID NO 6; 36pp; Japanese.
XX The invention relates to the novel testing method for diagnosing liver
CC cancer. The novel method comprises amplifying the amino terminal region
CC of a hepatitis-C virus NS3 gene using a predetermined primer and
CC determining the hepatitis-C virus in a base sequence of the obtained DNA
CC fragment. The novel testing method is useful for diagnosing liver cancer
CC and also used in a gene amplification technique, a clinical laboratory
CC test reagent, a polymerase chain reaction, a base sequence analysis and
CC genetic engineering. The method enables the detection of a hepatitis-C
CC virus having high carcinogenicity with high specificity. This sequence
CC represents the protein of the hepatitis-C virus NS3 gene of the
CC invention.
XX Sequence 3010 AA;

Query Match 95.2%; Score 3016; DB 7; Length 3010;
Best Local Similarity 96.3%; Pred. No. 1.5e-278;
Matches 569; Conservative 12; Mismatches 10; Indels 0; Gaps 0;

QY 5 SMSYTWTCALITPCAAERSKLPINALNSLRLHNLVYSTRSASLRQKKVTPDRLOVL 64
DB 2420 SMSYTWTCALITPCAAERSKLPINPLNSLRLHNSWVYSTRSASLRQKKVTPDRLOVL 2479
QY 65 DDHYRDVLKEMKAKASTVKAALLSVERACKLTPPHSAKSKFGYKADVRSLSSRAVNHIR 124
DB 2480 DDHYRDVLKEMKAKASTVKARLLSTEEACKLTPPHSAKSKFGYKADVRSLSSRAVNHIR 2539
QY 125 SVWKDLLEDTPPIOTTIMAKNEVFCVQPEKGRKPARLIVPPDLGVRVCEKMALYDVVS 184
DB 2540 SVWKDLLEDTPPIOTTIMAKNEVFCVQPEKGRKPARLIVPPDLGVRVCEKMALYDVVS 2599
QY 185 TLPOAVMGSSYGFQYSPKQRFVFLVNTWKAKCPMGFSYDTRCFDSTVTEINDIRVEESII 244
DB 2600 TLPOAVMGFSYGFQYSPQGRVFLVNTWKSKCPMGFSYDTRCFDSTVTEINDIRVEESII 2659
QY 245 QCCDLAPARQAIRSLRRLYVGGPMTNSKQNGCYRCRAGSVLTTSCGNTLTCLYKAA 304
DB 2660 QCCDLAPARQAIRSLRRLYVGGPLTNSKQNGCYRCRAGSVLTTSCGNTLTCLYKAT 2719
QY 305 AACRAAKLQDCTMLVNGDDLWVCESAGTQEDAAALRAFTAMTRYSPAGDPPOPEYDL 364
DB 2720 AACRAAKLQDCTMLVNGDDLWVCESAGTQEDAAALRAFTAMTRYSPAGDPPOPEYDL 2779
QY 365 ELITSCSNVSAHDAAGKRVYLLTRDPTVPLARAWEETARTHTPVNSWLGNIIMVAPTLW 424
DB 2780 ELITSCSNVSAHDAAGKRVYLLTRDPTVPLARAWEETARTHTPVNSWLGNIIMVAPTLW 2839
QY 425 ARMLMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPOIIRLHGLSAFSLHSYSPG 484

DB 2840 ARMLMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPOIIRLHGLSAFSLHSYSPG 2899
QY 485 EINRVASCLRLKLGVPPLRVWRHRSVRKALLSQCGRAAICGKLYFNWAVRKLKLTPIP 544
DB 2900 EINRVASCLRLKLGVPPLRVWRHRSVRKALLSQCGRAATCGKLYFNWAVRKLKLTPIP 2959
QY 545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMCLCLLLSVGVGYLLPNR 595
DB 2960 AASQLDLSGWFVAGYSGGDIYHSLSRARPRWFMCLCLLLSVGVGYLLPNR 3010

RESULT 9

AAW98022
ID AAW98022 standard; protein; 3010 AA.

XX AAW98022;

DT 21-JUN-1999 (first entry)

DE Infectious hepatitis C virus genotype 1b strain HC-J4 protein.

XX HCV; infectious clone; infection; diagnosis; therapy; vaccine; screening;

KW assay; antiviral; virucide.

XX Hepatitis C virus.

PN WO9904008-A2.

XX 28-JAN-1999.

PF 16-JUL-1998; 98WO-US014688.

PR 18-JUL-1997; 97US-0053062P.

PR 27-JAN-1998; 98US-00014416.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Yanagi M, Bukh J, Emerson SU, Purcell RH;

XX WPI; 1999-132252/11.

DR N-PSDB; AAX24843.

XX New isolated hepatitis C virus nucleic acids - used to develop products
PT for the diagnosis, prevention and treatment of HCV infections and for
PT developing screening assays.

XX Claim 2; Fig 14G-H; 126pp; English.

XX This protein is encoded by the infectious hepatitis C virus (HCV)
CC genotype 1b strain HC-J4 genome (see AAX24833). HC-J4 was obtained from
CC acute phase plasma of a chimpanzee infected with serum containing HC-
CC J4/91. The infectious nucleic acid sequence can be used to produce
CC chimeric genomes (see AAX24833) consisting of the open reading frames of
CC infectious nucleic acid sequences of other genotypes (including genotypes
CC 1-6) and subtypes (such as 1b, 2a, 2b, 2c, 3a, 4a-f, 5a and 6a) of HCV.
CC The invention also relates to the introduction of mutations or deletions
CC into infectious nucleic acid sequences in order to produce an attenuated
CC HCV virus suitable for vaccine development. Infectious nucleic acid
CC sequences can also be used to produce attenuated virus via passage in
CC vitro or in vivo of the viruses produced by transfection of a host cell
CC with the infectious nucleic acid sequence. Vaccines comprising one or
CC more polypeptides made from the infectious nucleic acid sequence are used
CC to immunise mammals, especially humans, against hepatitis C. The nucleic
CC acid sequences can also be used to induce protective immunity against the
CC virus. The nucleic acid sequences or their encoded proteases (e.g. NS3
CC protease) can additionally be used to develop screening assays to
CC identify antiviral agents for HCV

XX Sequence 3010 AA;

Query Match 95.2%; Score 3015; DB 2; Length 3010;
Best Local Similarity 95.9%; Pred. No. 1.9e-278;

Matches 567; Conservative 16; Mismatches 8; Indels 0; Gaps 0;

Qy 5 SMSYTWTCALITPCAAEESKLPINALSNLRLHNLVYTTSSRSASLRQKKVTPDRQLVL 64
 Db 2420 SMSYTWTCALITPCAAEESKLPINPLSNLRLHNLVYTTSSRSASLRQKKVTPDRQLVL 2479

Qy 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYCAKOVRSLSRAVNHIR 124
 Db 2480 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYCAKOVRSLSRAVNHIR 2539

Qy 125 SWKDLLEDTPDTIOTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKALYDVVS 184
 Db 2540 SWKDLLEDTPDTIOTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKALYDVVS 2599

Qy 185 TLPOAVMGSSYGFQYSPKQRFVFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
 Db 2600 TLPOAVMGSSYGFQYSPKQRFVFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 2659

Qy 245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGONCGYRRCRASGVLTTCGNTLTCLYKAA 304
 Db 2660 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGONCGYRRCRASGVLTTCGNTLTCLYKAT 2719

Qy 305 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
 Db 2720 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAALRAFTTEAMTRYSAAPPDPPQPEYDL 2779

Qy 365 ELITSCSSNVSAHDASGRVYLTTRDPTVPLARAAMETARHTPVNSWLNINMYAPTLW 424
 Db 2780 ELITSCSSNVSAHDASGRVYLTTRDPTVPLARAAMETARHTPVNSWLNINMYAPTLW 2839

Qy 425 ARMTLMTFFSILLAQOEKALDQCIYGACYSIEPLDLPOIIERLHGLSAFSLHSYSPG 484
 Db 2840 ARMTLMTFFSILLAQOEKALDQCIYGACYSIEPLDLPOIIERLHGLSAFTLHSYSPG 2899

Qy 485 EBNRVASCLRLKGLVPPPLRWHRARSVRKALLSOGGAAICGKYLFWAVRTKLTPIIP 544
 Db 2900 EBNRVASCLRLKGLVPPPLRWHRARSVRKALLSOGGAAICGKYLFWAVRTKLTPIIP 2959

Qy 545 AASRLDLSGMFWAGYSGGDIYHSLSRARPRWFMCLLLLSVGVGIYLLPNR 595
 Db 2960 AASRLDLSGMFWAGYSGGDIYHSLSRARPRWFMCLLLLSVGVGIYLLPNR 3010

RESULT 10
 AAB31170
 ID AAB31170 standard; protein; 3010 AA.
 XX AC AAB31170;
 XX 02-APR-2001 (first entry)
 DE Amino acid sequence of a hepatitis C virus (HCV) clone genotype 1b.
 XX KW Chimeric virus; bovine viral diarrhoea virus; BVDV; hepatitis C virus;
 XX KW HCV; vaccine; viral inhibitor; antiviral.
 OS Hepatitis C virus.
 XX FN WO200075352-A2.
 XX PD 14-DEC-2000.
 XX 02-JUN-2000; 2000WO-US015527.
 PR 04-JUN-1999; 99US-0137817P.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX PI Nam J, Bukh J, Emerson SU, Purcell RH;
 XX WPI; 2001-071081/08.
 DR N-PSDB; AAC86939.
 XX

PT New nucleic acid comprising a chimeric bovine viral diarrhoea virus genome
 PT in which the (non-)structural region has been replaced by hepatitis C
 PT virus (HCV) genome useful for treating or preventing HCV signs and
 PT symptoms.

XX Disclosure; Fig 4G-H; 97pp; English.

PS The specification describes a nucleic acid comprising a chimeric virus
 XX genome, specifically bovine viral diarrhoea virus (BVDV) genome in which
 CC the (non-)structural region has been replaced by the (non-)structural
 CC region of a hepatitis C virus (HCV) genome. The nucleic acids comprising
 CC the chimeric virus and the chimeric virus are useful for identifying cell
 CC lines capable of supporting the replication of these chimeric viruses, in
 CC screening for neutralizing antibodies to HCV of different genotypes, in
 CC the production of HCV-BVDV virions, for the development of inactivated or
 CC attenuated vaccines to prevent HCV-BVDV in a mammal, in studying the
 CC molecular properties of HCV indirectly in vitro, and in identifying
 CC inhibitors of viral enzyme activity which would be useful as antiviral
 CC agents. Formulations or compositions comprising the chimeric viruses may
 CC be used to treat or prevent the signs and symptoms of HCV. The present
 CC sequence is encoded by a HCV clone, which is used to construct chimeric
 CC nucleic acids of the invention

XX SQ Sequence 3010 AA;

Query Match 95.2%; Score 3015; DB 4; Length 3010;
 Best Local Similarity 95.9%; Pred. No. 1.9e-278;
 Matches 567; Conservative 16; Mismatches 8; Indels 0; Gaps 0;

Qy 5 SMSYTWTCALITPCAAEESKLPINALSNLRLHNLVYTTSSRSASLRQKKVTPDRQLVL 64
 Db 2420 SMSYTWTCALITPCAAEESKLPINPLSNLRLHNLVYTTSSRSASLRQKKVTPDRQLVL 2479

Qy 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYCAKOVRSLSRAVNHIR 124
 Db 2480 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYCAKOVRSLSRAVNHIR 2539

Qy 125 SWKDLLEDTPDTIOTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKALYDVVS 184
 Db 2540 SWKDLLEDTPDTIOTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKALYDVVS 2599

Qy 185 TLPOAVMGSSYGFQYSPKQRFVFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
 Db 2600 TLPOAVMGSSYGFQYSPKQRFVFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 2659

Qy 245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGONCGYRRCRASGVLTTCGNTLTCLYKAA 304
 Db 2660 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGONCGYRRCRASGVLTTCGNTLTCLYKAT 2719

Qy 305 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
 Db 2720 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAALRAFTTEAMTRYSAAPPDPPQPEYDL 2779

Qy 365 ELITSCSSNVSAHDASGRVYLTTRDPTVPLARAAMETARHTPVNSWLNINMYAPTLW 424
 Db 2780 ELITSCSSNVSAHDASGRVYLTTRDPTVPLARAAMETARHTPVNSWLNINMYAPTLW 2839

Qy 425 ARMTLMTFFSILLAQOEKALDQCIYGACYSIEPLDLPOIIERLHGLSAFSLHSYSPG 484
 Db 2840 ARMTLMTFFSILLAQOEKALDQCIYGACYSIEPLDLPOIIERLHGLSAFTLHSYSPG 2899

Qy 485 EBNRVASCLRLKGLVPPPLRWHRARSVRKALLSOGGAAICGKYLFWAVRTKLTPIIP 544
 Db 2900 EBNRVASCLRLKGLVPPPLRWHRARSVRKALLSOGGAAICGKYLFWAVRTKLTPIIP 2959

Qy 545 AASRLDLSGMFWAGYSGGDIYHSLSRARPRWFMCLLLLSVGVGIYLLPNR 595
 Db 2960 AASRLDLSGMFWAGYSGGDIYHSLSRARPRWFMCLLLLSVGVGIYLLPNR 3010

RESULT 11
 ADO36227
 ID ADO36227 standard; protein; 3010 AA.

XX AC ADO36227;
 XX DB 26-AUG-2004 (first entry)
 XX DE Hepatitis C virus (HCV) J4L6 wild-type polyprotein.
 XX KW hepatotropic; virucide; vaccine; gene therapy; vaccine;
 XX KW Hepatitis C virus; HCV; core protein; HCV infection; vaccination;
 XX KW polyprotein.
 XX OS Hepatitis C virus.
 XX PN WO2004046175-A1.
 XX PD 03-JUN-2004.
 XX PF 13-NOV-2003; 2003WO-EP012793.
 XX PR 15-NOV-2002; 2002GB-00026722.
 XX PA (GLAX) GLAXO GROUP LTD.
 XX PI Brett S, Hamblin PA, Ogilvie L;
 XX DR WPI; 2004-420613/39.
 XX DR N-ESDB; ADO36222.
 XX PT New Hepatitis C virus (HCV) vaccine having a polynucleotide that encodes
 XX PT the polypeptide sequences of the HCV core and at least one other HCV
 XX PT protein, for use in medicine, particularly for manufacturing a medicament
 XX PT for treating HCV.
 XX PS Disclosure; Fig 6; 78pp; English.
 XX CC The invention describes a polynucleotide vaccine comprising a
 XX CC polynucleotide sequence (S1) encoding the Hepatitis C virus (HCV) Core
 XX CC protein and at least 1 other HCV protein, and causes expression of the
 XX CC proteins in cells (in which (S1) has been mutated or positioned relative
 XX CC to the polynucleotide sequence encoding the other HCV protein, so that
 XX CC the negative effect of the Core protein on expression of the other HCV
 XX CC protein is reduced). Also described are: a method of preventing or
 XX CC treating an HCV infection in a mammal, comprising administering the
 XX CC vaccine cited above to a mammal; and a method of vaccination of an
 XX CC individual, comprising taking a polynucleotide vaccine as cited above,
 XX CC coating the polynucleotide onto gold beads and delivering the gold beads
 XX CC into the skin. HCV nucleic acids, polypeptides, host cells, vectors and
 XX CC antibodies used in the methods, are also disclosed. The polynucleotide
 XX CC vaccine is useful in the manufacture of a medicament for the treatment of
 XX CC HCV. This is the amino acid sequence of the wild type HCV polyprotein.
 XX SQ Sequence 3010 AA;

Query Match 95.2%; Score 3015; DB 8; Length 3010;
 Best Local Similarity 95.9%; Pred. No. 1.9e-278;
 Matches 567; Conservative 16; Mismatches 8; Indels 0; Gaps 0;

Qy 5 SMSYTWGALITPCAEESEKLPINALNSLRHNNLVSTTSRSASIRQKVTDFRLQVL 64
 Db 2420 SMSYTWGALITPCAEESEKLPINPLNSLRHNNLVSTTSRSASIRQKVTDFRLQVL 2479
 Qy 65 DDHYRDVLKEMKAKSTVKALLSVEAEKLTTPHSAKSKFGYKADVRSLSRAVNHIR 124
 Db 2480 DDHYRDVLKEMKAKSTVKALLSIEAEKLTTPHSAKSKFGYKADVRSLSRAVNHIR 2539
 Qy 125 SVWKDLLEDTPITQTTIMAKNEVFCVQPEKGRKPARLIIFPDLGVRVCCKMALYDVVS 184
 Db 2540 SVWEDLLEDTPITDITIMAKSEVFCVQPEKGRKPARLIIFPDLGVRVCCKMALYDVVS 2599
 Qy 185 TLPOAVNGSSYGFQYSPKQRFVFLVNTWKAKCPMGFSYDFRCFDSVTENDIRVESIY 244
 Db 2600 TLPOAVNGSSYGFQYSPKQRFVFLVNTWKSKCPMGFSYDFRCFDSVTENDIRVESIY 2659

Qy 245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKQNGCYRRCRAGVLTTCGNTLTCLYKAA 304
 Db 2660 QCCDLAPEARQAIRSLTERLYVGGPMTNSKQNGCYRRCRAGVLTTCGNTLTCLYKAT 2719
 Qy 305 AACRAAKLQDCTMLVNGDDLIVICESAGTQEDAAASLRVFTTEAMTRYSAAPPQPEYDL 364
 Db 2720 AACRAAKLQDCTMLVNGDDLIVICESAGTQEDAAALRAFTTEAMTRYSAAPPQPEYDL 2779
 Qy 365 ELITSCSSNVSAHDASGRVYVLTTRDPTVPLARAAMETARHTPVNSWLGNIMYAPTLM 424
 Db 2780 ELITSCSSNVSAHDASGRVYVLTTRDPTVPLARAAMETARHTPVNSWLGNIMYAPTLM 2839
 Qy 425 ARMILMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
 Db 2840 ARMILMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2899
 Qy 485 EINRVASCLRLKGVPPPLRVWRHARSVRAKLLSOGGAAICGKYLFWNAVTKLTLTIP 544
 Db 2900 EINRVASCLRLKGVPPPLRVWRHARSVRAKLLSOGGAAICGKYLFWNAVTKLTLTIP 2959
 Qy 545 AASRLDLGWFVAGYSGGDIYHLSRARPRWFMCLLILLSVGGIYLLPNR 595
 Db 2960 AASRLDLGWFVAGYSGGDIYHLSRARPRWFMCLLILLSVGGIYLLPNR 3010

RESULT 12
 ADO79401
 ID ADO79401 standard; protein; 3010 AA.
 AC ADO79401;
 DT 26-AUG-2004 (first entry)
 DE Hepatitis C virus J4L6 genome wild-type polyprotein.
 KW HCV; polyprotein; vaccine; DNA immunisation; hepatotropic; virucide.
 OS Hepatitis C virus.
 OS Synthetic.
 PN WO2004046176-A1.
 PD 03-JUN-2004.
 PF 13-NOV-2003; 2003WO-EP012830.
 PR 15-NOV-2002; 2002GB-00026722.
 PA (GLAX) GLAXO GROUP LTD.
 PI Brett S, Hamblin PA, Ogilvie L;
 DR WPI; 2004-420614/39.
 DR N-ESDB; ADO79396.
 PT New Hepatitis C virus (HCV) vaccine having a polynucleotide that encodes
 PT the polypeptide sequences of the HCV core, NS3, NS4B and NS5B proteins,
 PT for use in medicine, in particular for manufacturing a medicament for the
 PT treatment of HCV.
 PS Disclosure; Fig 6; 79pp; English.
 CC The present sequence is the hepatitis C virus (HCV) J4L6 genome wild-type
 CC polyprotein sequence. HCV DNA vaccines of the invention comprise a
 CC polynucleotide that encodes the HCV proteins Core, NS3, NS4B and NS5B,
 CC and does not encode the NS4A and/or NS5A proteins. Preferably, the codon
 CC usage of the polynucleotide sequence resembles that of highly expressed
 CC human genes. The polynucleotides may encode individual proteins or fusion
 CC proteins. Preferred fusions include double fusions between NS4B and NS5B
 CC and between Core and NS3. The vaccines are useful for the treatment or
 CC prevention of an HCV infection.
 XX Sequence 3010 AA;

Query Match 95.2%; Score 3015; DB 8; Length 3010;
 Best Local Similarity 95.9%; Pred. No. 1.9e-278;
 Matches 567; Conservative 16; Mismatches 8; Indels 0; Gaps 0;

Qy 5 SMSYTWGALITPCAAEESKLPINALNSLLRHNLVYSTTSRSASLRQKQKVTDFRLQVL 64
 Db 2420 SMSYTWGALITPCAAEESKLPINPLNSLLRHNLVYATTSRSASLRQKQKVTDFRLQVL 2479

Qy 65 DDHYRDVLKEMKAKASTVAKLLSVEEACKLTPPHSAKSKFGYGAQVRSLSRAVNHIR 124
 Db 2480 DDHYRDVLKEMKAKASTVAKLLSIEEACKLTPPHSAKSKFGYGAQVRSLSRAVNHIR 2539

Qy 125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGRKPARLIIVFPDLGVRCCKMALYDWS 184
 Db 2540 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGRKPARLIIVFPDLGVRCCKMALYDWS 2599

Qy 185 TLPQAVMGSSYGFQYSPKQRFVFLVNTWKAKKCPMGFSYDFRCFDSVTENDIRVESIY 244
 Db 2600 TLPQAVMGSSYGFQYSPKQRFVFLVNTWKSKCPMGFSYDFRCFDSVTENDIRVESIY 2659

Qy 245 QCCDLAPEARQAIRSLTERLYVGPMTNSKGQNGCYRRCRASGLVLTSCGNLTTCYLKAA 304
 Db 2660 QCCDLAPEARQAIRSLTERLYVGPMTNSKGQNGCYRRCRASGLVLTSCGNLTTCYLKAT 2719

Qy 305 AACRAAKLQDCTMLVNGDDLVIKESAGTQEDAAASLRVFTAMTRYSAAPGDPPEYDL 364
 Db 2720 AACRAAKLQDCTMLVNGDDLVIKESAGTQEDAAALRAFTAMTRYSAAPGDPPEYDL 2779

Qy 365 ELITSCSSNVSAHDASGRVYLTDPVTPPLARAAMETARHTPVNSGLNIIIMYAPTLW 424
 Db 2780 ELITSCSSNVSAHDASGRVYLTDPVTPPLARAAMETARHTPVNSGLNIIIMYAPTLW 2839

Qy 425 ARMTLMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPOIIRLHGLSAFSLHSYSPG 484
 Db 2840 ARMTLMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPOIIRLHGLSAFSLHSYSPG 2899

Qy 485 EINKVASCRLKGLVPPPLRVWRHRSVRKALLSOGGAAICGKYLEFNWAVTKLKTPI 544
 Db 2900 EINKVASCRLKGLVPPPLRVWRHRSVRKALLSOGGAAICGKYLEFNWAVTKLKTPI 2959

Qy 545 AASRLDLSGWFVAGYSGGDIYHLSLRARPRWFMCLLLLSVGVIYLLPNR 595
 Db 2960 AASRLDLSGWFVAGYSGGDIYHLSLRARPRWFMCLLLLSVGVIYLLPNR 3010

RESULT 13
 ID ADO36215
 AC ADO36215 standard; protein; 592 AA.
 AC ADO36215;
 DT 26-AUG-2004 (first entry)
 DE Hepatitis C virus (HCV) NS5B protein.
 KW hepatotropic; virucide; vaccine; gene therapy; vaccination; NS5B.
 KW Hepatitis C virus; HCV; core protein; HCV infection; vaccination; NS5B.
 OS Hepatitis C virus.
 PN W02004046175-A1.
 PD 03-JUN-2004.
 XX 13-NOV-2003; 2003WO-EP012793.
 XX 15-NOV-2002; 2002GB-00026722.
 XX (GLAX) GLAXO GROUP LTD.
 XX Brett S, Hamblin PA, Ogilvie L;

DR WPI; 2004-420613/39.
 DR N-PSDB; ADO36226.
 XX New Hepatitis C virus (HCV) vaccine having a polynucleotide that encodes
 PT the polypeptide sequences of the HCV core and at least one other HCV
 PT protein, for use in medicine, particularly for manufacturing a medicament
 PT for treating HCV.
 XX
 XX Example 2; Page 27; 78pp; English.

XX The invention describes a polynucleotide vaccine comprising a
 CC polynucleotide sequence (S1) encoding the Hepatitis C virus (HCV) Core
 CC protein and at least 1 other HCV protein, and causes expression of the
 CC proteins in cells (in which (S1) has been mutated or positioned relative
 CC to the polynucleotide sequence encoding the other HCV protein, so that
 CC the negative effect of the Core protein on expression of the other HCV
 CC protein is reduced). Also described are: a method of preventing or
 CC treating an HCV infection in a mammal, comprising administering the
 CC vaccine cited above to a mammal; and a method of vaccination of an
 CC individual, comprising taking a polynucleotide vaccine as cited above,
 CC coating the polynucleotide onto gold beads and delivering the gold beads
 CC into the skin. HCV nucleic acids, polypeptides, host cells, vectors and
 CC antibodies used in the methods, are also disclosed. The polynucleotide
 CC vaccine is useful in the manufacture of a medicament for the treatment of
 CC HCV. This is the amino acid sequence of the HCV NS5B protein.

XX
 SQ Sequence 592 AA;

Query Match 95.2%; Score 3014; DB 8; Length 592;
 Best Local Similarity 96.1%; Pred. No. 1.8e-279;
 Matches 569; Conservative 14; Mismatches 9; Indels 0; Gaps 0;

Qy 4 MSMSYTWGALITPCAAEESKLPINALNSLLRHNLVYSTTSRSASLRQKQKVTDFRLQV 63
 Db 1 MSMSYTWGALITPCAAEESKLPINPLNSLLRHNLVYATTSRSASLRQKQKVTDFRLQV 60

Qy 64 LDDHYRDVLKEMKAKASTVAKLLSVEEACKLTPPHSAKSKFGYGAQVRSLSRAVNHIR 123
 Db 61 LDDHYRDVLKEMKAKASTVAKLLSIEEACKLTPPHSAKSKFGYGAQVRSLSRAVNHIR 120

Qy 124 RSVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGRKPARLIIVFPDLGVRCCKMALYDVV 183
 Db 121 RSVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGRKPARLIIVFPDLGVRCCKMALYDVV 180

Qy 184 STLPQAVMGSSYGFQYSPKQRFVFLVNTWKAKKCPMGFSYDFRCFDSVTENDIRVESI 243
 Db 181 STLPQAVMGSSYGFQYSPKQRFVFLVNTWKSKCPMGFSYDFRCFDSVTENDIRVESI 240

Qy 244 YQCCDLAPEARQAIRSLTERLYVGPMTNSKGQNGCYRRCRASGLVLTSCGNLTTCYLKA 303
 Db 241 YQCCDLAPEARQAIRSLTERLYVGPMTNSKGQNGCYRRCRASGLVLTSCGNLTTCYLKA 300

Qy 304 AAACRAAKLQDCTMLVNGDDLVIKESAGTQEDAAASLRVFTAMTRYSAAPGDPPEYD 363
 Db 301 AAACRAAKLQDCTMLVNGDDLVIKESAGTQEDAAALRAFTAMTRYSAAPGDPPEYD 360

Qy 364 LELITSCSSNVSAHDASGRVYLTDPVTPPLARAAMETARHTPVNSGLNIIIMYAPTL 423
 Db 361 LELITSCSSNVSAHDASGRVYLTDPVTPPLARAAMETARHTPVNSGLNIIIMYAPTL 420

Qy 424 WARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPOIIRLHGLSAFSLHSYSP 483
 Db 421 WARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPOIIRLHGLSAFSLHSYSP 480

Qy 484 GEINRVASCLRLKGLVPPPLRVWRHRSVRKALLSOGGAAICGKYLEFNWAVTKLKTPI 543
 Db 481 GEINRVASCLRLKGLVPPPLRVWRHRSVRKALLSOGGAAICGKYLEFNWAVTKLKTPI 540

Qy 544 PAASRLDLSGWFVAGYSGGDIYHLSLRARPRWFMCLLLLSVGVIYLLPNR 595
 Db 541 PAASRLDLSGWFVAGYSGGDIYHLSLRARPRWFMCLLLLSVGVIYLLPNR 592

RESULT 14

ADO79389
ID ADO79389 standard; protein; 592 AA.

XX ADO79389;

DT 26-AUG-2004 (first entry)

DE Hepatitis C virus NS5B protein.

XX HCV; NS5B; vaccine; DNA immunisation; hepatotropic; virucide; mutant;

KW mtein.

XX Hepatitis C virus.

OS Synthetic.

XX WO2004046176-A1.

XX 03-JUN-2004.

XX 13-NOV-2003; 2003WO-BP012830.

XX 15-NOV-2002; 2002GB-00026722.

XX (GLAXO) GLAXO GROUP LTD.

XX Brett S, Hamblin PA, Ogilvie L;

XX WPI; 2004-420614/39.

XX N-PSDB; ADO79400.

XX New Hepatitis C virus (HCV) vaccine having a polynucleotide that encodes

PT the polypeptide sequences of the HCV core, NS3, NS4B and NS5B proteins,

PT for use in medicine, in particular for manufacturing a medicament for the

PT treatment of HCV.

XX Example 2; Page 27; 79pp; English.

XX The present sequence is that of the NS5B protein of hepatitis C virus

CC (HCV). The sequence is the translation sequence of a polynucleotide in

CC which codon usage was altered to resemble that of highly expressed human

CC genes. HCV vaccines of the invention comprise a polynucleotide that

CC encodes the HCV proteins Core, NS3, NS4B and NS5B, and does not encode

CC the NS4A and/or NS5A proteins. The proteins may be expressed as

CC individual proteins or as fusion proteins. Preferred fusions include

CC double fusions between NS4B and NS5B and between Core and NS3. The

CC vaccines are useful for the treatment or prevention of an HCV infection.

XX SQ Sequence 592 AA;

Query Match 95.2%; Score 3014; DB 8; Length 592;

Best Local Similarity 96.1%; Pred. No. 1.8e-279;

Matches 569; Conservative 14; Mismatches 9; Indels 0; Gaps 0;

Qy 4 MSMSYTWGALITPCAAEESKLPINALNSLRHNLVYSTTSASLRQKKVTFDRLOV 63

Db 1 MSMSYTWGALITPCAAEESKLPINPLNSLRHNLVYATTSASLRQKKVTFDRLOV 60

Qy 64 LDDHYRDVLEKMAKASTVKAISVEACKLTPPHSAKSKFGYCAKDVRSLSRAVNH 123

Db 61 LDDHYRDVLEKMAKASTVKAISVEACKLTPPHSAKSKFGYCAKDVRSLSRAVNH 120

Qy 124 RSVWKDLLEDTDPTIOTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMAKYDVV 183

Db 121 RSVWKDLLEDTDPTIOTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMAKYDVV 180

Qy 184 STLPAQVMGSSYFGYSPKQRFELVNTWKAICPMGFSYTRCFDSTVTENDIRVEESI 243

Db 181 STLPAQVMGSSYFGYSPKQRFELVNTWKAICPMGFSYTRCFDSTVTENDIRVEESI 240

Qy 244 YQCCDLAPARQAIRSLTERLYVGGPMTNSKQNGCYRRRCRASGVLITSCGNLTCTYLKA 303

Db 241 YQCCDLAPARQAIRSLTERLYVGGPMTNSKQNGCYRRRCRASGVLITSCGNLTCTYLKA 300

Qy 304 AACRAAKLQDCTMLVNGDDLVIWICESAGTQEDAAASLRVFTTEAMTRYAPPDGPQPEYD 363

Db 301 TAACRAAKLQDCTMLVNGDDLVIWICESAGTQEDAAALRAFTTEAMTRYAPPDGPQPEYD 360

Qy 364 LELITSCSSNVSAHDASGKRVYVLTTRDPTVPLARAANETARHTTVNSWLGNIIMYAPT 423

Db 361 LELITSCSSNVSAHDASGKRVYVLTTRDPTVPLARAANETARHTTVNSWLGNIIMYAPT 420

Qy 424 WARMLMTHFFSILLAQLEKALDCQIYGACYSIEPLDLPOIIERLHGLSAFSLHSYSP 483

Db 421 WARMLMTHFFSILLAQLEKALDCQIYGACYSIEPLDLPOIIERLHGLSAFSLHSYSP 480

Qy 484 GEINRVASCLRLKGVPPPLRVWRHARSVRKLLSQGAAICGKYLFWNAVTKLKLTP 543

Db 481 GEINRVASCLRLKGVPPPLRVWRHARSVRKLLSQGAAICGKYLFWNAVTKLKLTP 540

Qy 544 PAASRLDLSGWFGVAGYSGGDIYHSLSRARPMFPLCLLLLSVGVIYLLPNR 595

Db 541 PAASRLDLSGWFGVAGYSGGDIYHSLSRARPMFPLCLLLLSVGVIYLLPNR 592

RESULT 15

AA06423
ID AA06423 standard; protein; 3010 AA.

XX AA06423;

XX 20-MAR-2003 (revised)

DT 27-SEP-1999 (first entry)

XX

DE Non-A, non-B hepatitis virus polypeptide.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Search completed: September 22, 2005, 14:47:11
Job time : 172 secs

```

XX WPI: 1999-407152/35.
DR N-PSDB; AAX59394.
XX
PT New hepatitis virus polypeptides, useful for diagnosing and treating
PT hepatitis infections.
XX
PS Claim 2; Fig 2(1)-(16); 56pp; English.
XX
CC This sequence represents the non-A, non-B hepatitis virus (NANBV)
CC polypeptide, as predicted from cDNA (see AAX59394) containing the entire
CC open reading frame of the NANBV genome. To obtain this cDNA, NANBV RNAs
CC were extracted directly from NANBV particles contained in whole blood of
CC a patient having NANBV hepatitis, or from a resected liver of a patient
CC having NANBV hepatitis and liver cancer. The RNA was then converted to
CC double-stranded cDNA. A cDNA library was produced and screened using
CC serum from a convalescent patient having acute NANBV hepatitis and serum
CC from a patient having chronic NANBV hepatitis. The isolated cDNA allows
CC recombinant production of NANBV antigen polypeptides in microbial or
CC eukaryotic cell culture. The method provides the safe production of NANBV
CC antigens with high purity on a large scale at low cost without the
CC biohazard associated with multiplying virus in animals. Claimed NANBV
CC nucleotide sequences are useful for the recombinant production of
CC polypeptides useful as antigens for vaccines, and as diagnostic reagents.
CC (Updated on 20-MAR-2003 to correct PF field.) (Updated on 20-MAR-2003 to
CC correct PR field.)
XX
SQ Sequence 3010 AA;
    Query Match          95.2%; Score 3014; DB 2; Length 3010;
    Best Local Similarity 96.6%; Pred. No. 2,3e-278;
    Matches 571; Conservative 10; Mismatches 10; Indels 0; Gaps 0;
QY 5 SMSYTWTCALITPCAAEESKLPINALSNSLLRHNLVYSTTSRSASLRQKKVTFDRLOVL 64
Db 2420 SMSYTWTCALITPCAAEESKLPINALSNSLLRHNLVYATTSRSAGLRQKKVTFDRLOVL 2479
QY 65 DDHYRDVLKEMKASTYKALLSVEEACKLTPPHSAKSKFGYGNKDVYSSRAVNHIR 124
Db 2480 DDHYRDVLKEMKASTYKALLSVEEACKLTPPHSAKSKFGYGNKDVYSSRAVNHIR 2539
QY 125 SVWKDLLEDTDTPITIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMAKYDVVS 184
Db 2540 SVWKDLLEDTDTPITIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMAKYDVVS 2599
QY 185 TLPQVMGSSYGFQYSPQGVFEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
Db 2600 TLPQVMGSSYGFQYSPQGVFEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 2659
QY 245 QCCDLAPARQAIRSLTERLYVGGPMTNSKGNGYRCRACSGVLTSCGNTLTCYKAA 304
Db 2660 QCCDLAPARQAIRSLTERLYVGGPMTNSKGNGYRCRACSGVLTSCGNTLTCYKAA 2719
QY 305 AACRAKLQDCTMLVNGDDLVVICESAGTQEDASLRVFTAMTRYSAAPPDPOPEYDL 364
Db 2720 AACRAKLQDCTMLVNGDDLVVICESAGTQEDASLRVFTAMTRYSAAPPDPOPEYDL 2779
QY 365 ELITSCSSNVSAHDASGRVYVLTFRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTLW 424
Db 2780 ELITSCSSNVSAHDASGRVYVLTFRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTLW 2839
QY 425 ARMLIMTHFFSILLAOEQLKALDCQIYGACYSIEPLDLPQIIRLHGLSFAFLSHSPG 484
Db 2840 ARMLIMTHFFSILLAOEQLKALDCQIYGACYSIEPLDLPQIIRLHGLSFAFLSHSPG 2899
QY 485 EINRVASCLRLKLGVPPLRVWRHRSVRKALLSQGAAICGKYLFWNAVTKLKTPIIP 544
Db 2900 EINRVASCLRLKLGVPPLRVWRHRSVRKALLSQGAAICGKYLFWNAVTKLKTPIIP 2959
QY 545 AASRLDLSGWVFVAGYSGGDIYHLSLRARPRWFMCLCLLLSVGVGIYLLPNR 595
Db 2960 AASRLDLSGWVFVAGYSGGDIYHLSLRARPRWFMCLCLLLSVGVGIYLLPNR 3010

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2005, 14:40:28 ; Search time 19 Seconds
(without alignments)
3043.488 Million cell updates/sec

Title: US-10-712-479-2

Perfect score: 3167

Sequence: 1 MASMSMTWTGALITPCA.....LLSVGVGIYLLPNRHHHHH 601

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3016	95.2	3010	1 GNMVJC	genome polyprotein
2	3016	95.2	3010	1 GNMVTC	genome polyprotein
3	3003	94.8	3010	1 A45573	genome polyprotein
4	2968	93.7	3010	1 GNMVTV	genome polyprotein
5	2944	93.0	3010	1 S18030	genome polyprotein
6	2768	88.3	3011	1 S40770	genome polyprotein
7	2764	87.3	3011	1 GNMVVC	genome polyprotein
8	2725	86.0	3011	1 GNMVCH	genome polyprotein
9	2511	79.3	3014	1 JC5620	genome polyprotein
10	2392	75.5	3033	1 GNMVJ8	genome polyprotein
11	2380	75.1	3033	1 JQ1303	genome polyprotein
12	1886	59.6	365	2 JQ0879	NS5 protein - hepa
13	1750	55.3	365	2 JQ0880	NS5 protein - hepa
14	1471.5	46.5	874	2 JQ0883	genome polyprotein
15	1442	45.5	874	2 JQ0881	genome polyprotein
16	951	30.0	189	2 S60587	non-structural pro
17	774	24.4	284	2 PS0104	genome polyprotein
18	744	23.5	259	2 PS0102	genome polyprotein
19	719	22.7	156	2 D39109	genome polyprotein
20	667	21.1	135	2 S44214	genome polyprotein
21	654	20.7	3005	2 T08841	polyprotein - dour
22	636	20.1	1435	2 T01075	polyprotein - hepa
23	629	19.9	2970	2 T08839	polyprotein - marm
24	569	18.0	113	2 PC1278	NS5 protein - hepa
25	563	17.8	113	2 PC1277	NS5 protein - hepa
26	562	17.7	113	2 PC1274	NS5 protein - hepa
27	562	17.7	113	2 PC1275	NS5 protein - hepa
28	560	17.7	113	2 PC1276	NS5 protein - hepa
29	513	16.2	113	2 PC1279	NS5 protein - hepa

RESULT 1

GNMVJC

genome polyprotein - hepatitis C virus (strain J)

N:Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5

C:Species: hepatitis C virus

C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004

C:Accession: A39253; PS0086

R:Kato, N.; Hijikata, M.; Ootsuyama, Y.; Nakagawa, M.; Ohkoshi, S.; Sugimura, T.; Shimoto Proc. Natl. Acad. Sci. U.S.A. 87, 9524-9528, 1990

A:Title: Molecular cloning of the human hepatitis C virus genome from Japanese patients v A:Reference number: A39253; MUID:91088550; PMID:2175903

A:Accession: A39253

A:Molecule type: genomic RNA

A:Residues: 1-3010 <Kat>

A:Cross-references: UNIPROT:P26662; GB:D90208; NID:G221610; PID:BA414233.1; PID:G221611 R:Kato, N.; Ohkoshi, S.; Shimotohno, K.

Proc. Jpn. Acad. 65B, 219-223, 1989

A:Title: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence variat A:Reference number: PS0085

A:Accession: PS0086

A:Molecule type: Genomic RNA

A:Residues: 2650-2707 <KA2>

A:Experimental source: Japanese isolate

C:Comment: The cleavage sites of this polyprotein have not been determined.

C:Superfamily: hepatitis C virus genome polyprotein

C:Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serine F:2-115/Product: capsid protein C #status predicted <CPC>

F:116-191/Product: envelope protein M #status predicted <EPM>

F:192-389/Product: major envelope protein E #status predicted <ME>

F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>

F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>

F:1007-1615/Product: hepatitisin #status predicted <NS3>

F:1230-1237/Region: nucleotide-binding motif A (P-loop)

F:1312-1317/Region: nucleotide-binding motif B

F:1316-1319/Region: DEXH motif

F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>

F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>

F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>

F:196,209,234,250,305,325,417,423,430,448,532,556,576,623,645,1213,1255,2041,2077,2240,2;

Query Match 95.2%; Score 3016; DB 1; Length 3010;

Best Local Similarity 96.3%; Pred. No. 3.8e-223;

Matches 569; Conservative 12; Mismatches 10; Indels 0; Gaps 0;

Qy 5 SMSYTWGALITPCAABESKLPINALSNLRHNLVYSTTSRSASLRQKKVTFDLQVL 64

Db 2420 SMSYTWGALITPCAABESKLPINPLSNLRHNSMVYSTTSRSASLRQKKVTFDLQVL 2479

Qy 65 DDHYRDVLKEMKAKASTVAKLLSVBEACKLTTPPHSAKSKFGYKGVRSLSRAVNHIR 124

Db 2480 DDHYRDVLKEMKAKASTVAKLLSIEACKLTTPPHSAKSKFGYKGVRSLSRAVNHIR 2539

Qy 65 DDHYRDVLKEMKAKASTVAKLLSVBEACKLTTPPHSAKSKFGYKGVRSLSRAVNHIR 124

Db 2480 DDHYRDVLKEMKAKASTVAKLLSIEACKLTTPPHSAKSKFGYKGVRSLSRAVNHIR 2539

Best Local Similarity 95.9%; Pred. No. 3.8e-222; Matches 567; Conservative 15; Mismatches 9; Indels 0; Gaps 0;

QY 5 SMSYTWTCALITPCAAEESKLPINALSNLSLRHNLVSTTSRSASLRQKVKVTFORLOVL 64
DB 2420 SMSYTWTCALITPCAAEESKLPINALSNLSLRHNLVSTTSRSASLRQKVKVTFORLOVL 2479

QY 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
DB 2480 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 2539

QY 125 SVWKDLLEDTPITOTIMAKNEVFCVQPEKGRKPARLIVPPDLGVRVCEKQALYDVVS 184
DB 2540 SVWKDLLEDTPITOTIMAKNEVFCVQPEKGRKPARLIVPPDLGVRVCEKQALYDVVS 2599

QY 185 TLPOAVMGSSYGFQYSPQORVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESY 244
DB 2600 TLPOAVMGSSYGFQYSPQORVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESY 2659

QY 245 QCCDLAPPEARQAIRSLTERLYVGGPMTNSKGQNCYRRCRASGLVLTSCGNTLTCTYLKAA 304
DB 2660 QCCDLAPPEARQAIRSLTERLYVGGPMTNSKGQNCYRRCRASGLVLTSCGNTLTCTYLKAT 2719

QY 305 AACRAAKLQDCTMLVNGDDLVI CESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
DB 2720 AACRAAKLQDCTMLVNGDDLVI CESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 2779

QY 365 ELITSCSSNVSAHDASGRVYVLTTRDPTVPLARAWEATARTPTVNSWLGNIIMTAPTLM 424
DB 2780 ELITSCSSNVSAHDASGRVYVLTTRDPTVPLARAWEATARTPTVNSWLGNIIMTAPTLM 2839

QY 425 ARMILMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLQPIIERLHGLSAFSLHSYSPG 484
DB 2840 ARMILMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLQPIIERLHGLSAFSLHSYSPG 2899

QY 485 EIRNVASCLRLKLGVPPLRWRHRSVRKALLSQGGRRAICGKYLFWNAVRTKLTPTIP 544
DB 2900 EIRNVASCLRLKLGVPPLRWRHRSVRKALLSQGGRRAICGKYLFWNAVRTKLTPTIP 2959

QY 545 AASRLDLSGFWFVAGYSGGDIYHLSLRARPRWFMCLCLLLSVGVGYLLLPNR 595
DB 2960 AASRLDLSGFWFVAGYSGGDIYHLSLRARPRWFMCLCLLLSVGVGYLLLPNR 3010

RESULT 4
GNWVTW
genome polyprotein - hepatitis C virus (strain Taiwan)
N:Contains: capsid protein C; envelope protein M; hepatitis C virus (strain Taiwan)
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C:Accession: A40244
R:Chen, P.J.; Lin, M.H.; Tai, K.P.; Lin, P.C.; Lin, C.J.; Chen, D.S.
Virology 188, 102-113, 1992
A:Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the
A:Reference number: A40244; MUID:92230206; PMID:1314449
A:Accession: A40244
A:Molecule type: genomic RNA
A:Residues: 1-3010 <CHE>
A:Cross-references: UNIPROT:P29846; GB:M84754
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural
F:11-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <EPW>
F:192-389/Product: major envelope protein E #status predicted <MEE>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus genome polyprotein NS3 #status predicted <NS3>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>

F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,233,250,305,325,417,423,430,448,556,576,623,645,1213,1255,2041,207;
Query Match 93.7%; Score 2968; DB 1; Length 3010;
Best Local Similarity 94.9%; Pred. No. 1.9e-219; Matches 561; Conservative 15; Mismatches 15; Indels 0; Gaps 0;

QY 5 SMSYTWTCALITPCAAEESKLPINALSNLSLRHNLVSTTSRSASLRQKVKVTFORLOVL 64
DB 2420 SMSYTWTCALITPCAAEESKLPINALSNLSLRHNLVSTTSRSASLRQKVKVTFORLOVL 2479

QY 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
DB 2480 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 2539

QY 125 SVWKDLLEDTPITOTIMAKNEVFCVQPEKGRKPARLIVPPDLGVRVCEKQALYDVVS 184
DB 2540 SVWKDLLEDTPITOTIMAKNEVFCVQPEKGRKPARLIVPPDLGVRVCEKQALYDVVS 2599

QY 185 TLPOAVMGSSYGFQYSPQORVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESY 244
DB 2600 TLPOAVMGSSYGFQYSPQORVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESY 2659

QY 245 QCCDLAPPEARQAIRSLTERLYVGGPMTNSKGQNCYRRCRASGLVLTSCGNTLTCTYLKAA 304
DB 2660 QCCDLAPPEARQAIRSLTERLYVGGPMTNSKGQNCYRRCRASGLVLTSCGNTLTCTYLKAS 2719

QY 305 AACRAAKLQDCTMLVNGDDLVI CESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
DB 2720 AACRAAKLQDCTMLVNGDDLVI CESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 2779

QY 365 ELITSCSSNVSAHDASGRVYVLTTRDPTVPLARAWEATARTPTVNSWLGNIIMTAPTLM 424
DB 2780 ELITSCSSNVSAHDASGRVYVLTTRDPTVPLARAWEATARTPTVNSWLGNIIMTAPTLM 2839

QY 425 ARMILMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLQPIIERLHGLSAFSLHSYSPG 484
DB 2840 ARMILMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLQPIIERLHGLSAFSLHSYSPG 2899

QY 485 EIRNVASCLRLKLGVPPLRWRHRSVRKALLSQGGRRAICGKYLFWNAVRTKLTPTIP 544
DB 2900 EIRNVASCLRLKLGVPPLRWRHRSVRKALLSQGGRRAICGKYLFWNAVRTKLTPTIP 2959

QY 545 AASRLDLSGFWFVAGYSGGDIYHLSLRARPRWFMCLCLLLSVGVGYLLLPNR 595
DB 2960 AASRLDLSGFWFVAGYSGGDIYHLSLRARPRWFMCLCLLLSVGVGYLLLPNR 3010

RESULT 5
S18030
genome polyprotein - hepatitis C virus (isolate JKI)
N:Contains: capsid protein C; envelope protein M; hepatitis C virus (isolate JKI)
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
A:Variety: isolate JKI
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C:Accession: S18030; S33570; A48332; S18029
R:Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
A:Description: A whole genome of hepatitis C virus cDNA was isolated from a single patient
A:Reference number: S18028
A:Accession: S18030
A:Molecule type: genomic RNA
A:Residues: 1-3010 <HON>
A:Cross-references: UNIPROT:Q69949; EMBL:X61596; NID:g59478; PIDN:CAA43793.1; PID:g59479
A:Experimental source: isolate JKI from an individual
R:Honda, M.; Kaneko, S.; Unoura, M.; Kobayashi, K.; Murakami, S.
Arch. Virol. 128, 163-169, 1993
A:Title: Sequence analysis of putative structural regions of hepatitis C virus isolated from
A:Reference number: A48332; MUID:93119270; PMID:8380322
A:Accession: S33570
A:Molecule type: genomic RNA

Db	2661	QCCLDLPQARVAIKSLTERLYVGGPLTNSRGENGCGYRRCRASGLVLTSCGNLTFCYIKAR	2720
Qy	305	AACRAAKLQDCTMLVNGDDLVLVICSAGTQEDAAASLRVFTEAMTRYSAAPGDPPOPEYDL	364
Db	2721	AACRAAGLQDCTMLVCGDDLVLVICSAGVQEDAAASLRAFTTEAMTRYSAAPGDPPOPEYDL	2780
Qy	365	ELITSCSSNVSAHDASGKRVYLLTRDPTVPLARAAWETARHTPVSNLGNIIMYAPTLW	424
Db	2781	ELITSCSSNVSAHDGAGKRVYLLTRDPTTPLARAAWETARHTPVSNLGNIIMFAPTLW	2840
Qy	425	ARMILMTHTFFSLTAQEOLEKALDCQIYCACYSIEPLDLPOIITLRLHGLSAFSLHSYSPG	484
Db	2841	ARMILMTHTFFSVLIARQLEQALDCIYACYSIEPLDLPIIQLRLHGLSAFSLHSYSPG	2900
Qy	485	EINRVASCLRLGVPPFLRVRRHARSVRAKLLSQGAAICGKYLFWAVRTKLKLTPIP	544
Db	2901	EINRVASCLRLGVPPFLRAWHARSVRARLLARGGAAICGKYLFWAVRTKLKLTPIA	2960
Qy	545	AASRLDLSGFWAGVSGGDIYHSLSRARPRFWMCLLLLSVGVGVIYLLPNR	595
Db	2961	AAGQLDLSGFWTAGYSGGDIYHSVSHARPRWTFCLLLLAAGVGIYLLPNR	3011
RESULT 8			
GNWVCH			
Genome polypeptide - hepatitis C virus (strain H)			
N:Contains: capsid protein C; envelope protein M; hepatitis virus (BC 3.4.21.98) (nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5			
C:Species: hepatitis C virus			
A:Note: host Homo sapiens (man)			
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004			
C:Accession: A36814; A41546			
R:Inchauspe, G.; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M. submitted to GenBank, July 1992			
A:Description: Genomic structure of the human prototype strain H of hepatitis C virus: comparison			
A:Reference number: A36814			
A:Accession: A36814			
A:Molecule type: genomic RNA			
A:Residues: 1-3011 <NC>			
A:Cross-references: UNIPROT:P27958; GB:M67463; NID:G329737; PIDN:AAA45534.1; PID:G329738			
R:Inchauspe, G.; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M. Proc. Natl. Acad. Sci. U.S.A. 88, 10292-10296, 1991			
A:Title: Genomic structure of the human prototype strain H of hepatitis C virus: comparison			
A:Reference number: A41546; MUID:92052256; PMID:1658800			
A:Contents: annotation			
A:Note: neither amino acid nor nucleotide sequence is given			
C:Superfamily: hepatitis C virus genome polypeptide			
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural			
F:1-115/Product: capsid protein C #status predicted <CPC>			
F:116-191/Product: envelope protein M #status predicted <EPM>			
F:192-389/Product: major envelope protein E #status predicted <MEE>			
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>			
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>			
F:1007-1615/Product: hepatitis virus #status predicted <NS3>			
F:1230-1237/Region: nucleotide-binding motif A (P-loop)			
F:1312-1317/Region: nucleotide-binding motif B			
F:1316-1319/Region: DEXH motif			
F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4a>			
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4b>			
F:2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>			
F:1396, 209, 234, 305, 325, 417, 423, 430, 448, 476, 532, 540, 556, 576, 623, 645, 1213, 1255, 2041, 2240, 2340, 2341, 2342, 2343, 2344, 2345, 2346, 2347, 2348, 2349, 2350, 2351, 2352, 2353, 2354, 2355, 2356, 2357, 2358, 2359, 2360, 2361, 2362, 2363, 2364, 2365, 2366, 2367, 2368, 2369, 2370, 2371, 2372, 2373, 2374, 2375, 2376, 2377, 2378, 2379, 2380, 2381, 2382, 2383, 2384, 2385, 2386, 2387, 2388, 2389, 2390, 2391, 2392, 2393, 2394, 2395, 2396, 2397, 2398, 2399, 2400, 2401, 2402, 2403, 2404, 2405, 2406, 2407, 2408, 2409, 2410, 2411, 2412, 2413, 2414, 2415, 2416, 2417, 2418, 2419, 2420, 2421, 2422, 2423, 2424, 2425, 2426, 2427, 2428, 2429, 2430, 2431, 2432, 2433, 2434, 2435, 2436, 2437, 2438, 2439, 2440, 2441, 2442, 2443, 2444, 2445, 2446, 2447, 2448, 2449, 2450, 2451, 2452, 2453, 2454, 2455, 2456, 2457, 2458, 2459, 2460, 2461, 2462, 2463, 2464, 2465, 2466, 2467, 2468, 2469, 2470, 2471, 2472, 2473, 2474, 2475, 2476, 2477, 2478, 2479, 2480, 2481, 2482, 2483, 2484, 2485, 2486, 2487, 2488, 2489, 2490, 2491, 2492, 2493, 2494, 2495, 2496, 2497, 2498, 2499, 2500, 2501, 2502, 2503, 2504, 2505, 2506, 2507, 2508, 2509, 2510, 2511, 2512, 2513, 2514, 2515, 2516, 2517, 2518, 2519, 2520, 2521, 2522, 2523, 2524, 2525, 2526, 2527, 2528, 2529, 2530, 2531, 2532, 2533, 2534, 2535, 2536, 2537, 2538, 2539, 2540, 2541, 2542, 2543, 2544, 2545, 2546, 2547, 2548, 2549, 2550, 2551, 2552, 2553, 2554, 2555, 2556, 2557, 2558, 2559, 2560, 2561, 2562, 2563, 2564, 2565, 2566, 2567, 2568, 2569, 2570, 2571, 2572, 2573, 2574, 2575, 2576, 2577, 2578, 2579, 2580, 2581, 2582, 2583, 2584, 2585, 2586, 2587, 2588, 2589, 2590, 2591, 2592, 2593, 2594, 2595, 2596, 2597, 2598, 2599, 2600, 2601, 2602, 2603, 2604, 2605, 2606, 2607, 2608, 2609, 2610, 2611, 2612, 2613, 2614, 2615, 2616, 2617, 2618, 2619, 2620, 2621, 2622, 2623, 2624, 2625, 2626, 2627, 2628, 2629, 2630, 2631, 2632, 2633, 2634, 2635, 2636, 2637, 2638, 2639, 2640, 2641, 2642, 2643, 2644, 2645, 2646, 2647, 2648, 2649, 2650, 2651, 2652, 2653, 2654, 2655, 2656, 2657, 2658, 2659, 2660, 2661, 2662, 2663, 2664, 2665, 2666, 2667, 2668, 2669, 2670, 2671, 26			

```

QY 125 SVWKDLLEDDTPTIOTTIMAKNEVFCVQPEKGGKRPARLIVFPDLGVRVCEKMALYDVVS 184
DB 2541 SVWKDLLEDSVTPIDTTIMAKNEVFCVQPEKGGKRPARLIVFPDLGVRVCEKMALYDVVS 2600

QY 185 TLPQAVMGSSYGFQYSPQORVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
DB 2601 KLPLAVMGSSYGFQYSPQORVEFLVQAKSKTTPMGLSYDTRCFDSTVTESDIRTEEAIY 2660

QY 245 QCCDLAPARQAIRSLTERLYVGGPMTNKSGQNGCYRRCRASGVLTTCGNTLTCYLKAA 304
DB 2661 QCCDLDPQARVAIKSLTERLYVGGPLTNSRGNGCYRRCRASRVLTTCGNTLTRYIKAR 2720

QY 305 AACRAAKLQDCTMLVNGDDLVIICESAGTQDDAASLRVFTAMTRYSAAPPDPPQPEYDL 364
DB 2721 AACRAAGLQDCTMLVCGDDLVIICESAGVQDDAASLRFTAMTRYSAAPPDPPQPEYDL 2780

QY 365 ELITSCSSNVSAHDGAKRVYVLTDRPTVPLARAAMETARHTPVNSWLGNIIMVAPTLW 424
DB 2781 ELITSCSSNVSAHDGAKRVYVLTDRPTVPLARAAMETARHTPVNSWLGNIIMVAPTLW 2840

QY 425 ARMLTMTHFFSILLAQEOLEKALDCQIYGACYSIEPLDLPQIIRLHGLSAFSLHSYSPG 484
DB 2841 ARMLTMTHFFSVLIARDQLEQALANCEIYGACYSIEPLDLPPIIQRLHGLSAFSLHSYSPG 2900

QY 485 EINRVASCLRLKGLVPPPLRVWHRARSVRAKLLSGGRAAICGKYLFWAVRTKLTPIIP 544
DB 2901 EINRVAACLRKGLVPPPLRAWRHRAWSVRARLLARGGKAAICGKYLFWAVRTKLTPIIT 2960

QY 545 AASRLDLSGFWAGYSGGDIYHSLSRARPRWFMCLCLLLSVGVGIYLLPNR 595
DB 2961 AAGRLDLSGFWTAGYSGGDIYHSHVSHARPRWFMFCLLLLAAGVGIYLLPNR 3011

RESULT 9
JC5620
genome polyprotein - hepatitis C virus (isolate EUH1480)
N;Contains: capsid protein C; envelope protein M; hepatitis C virus genome polyprotein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C;Accession: JC5620
R;Chamberlain, R.W.; Adams, N.J.; Taylor, L.A.; Simmonds, P.; Elliott, R.M.
Biochem. Biophys. Res. Commun. 236, 44-49, 1997
A;Title: The complete coding sequence of hepatitis C virus genotype 5a, the predominant
A;Reference number: JC5620; MUID:97366593; PMID:9223423
A;Accession: JC5620
A;Molecule type: mRNA
A;Residues: 1-3014 <CHA>
A;Cross-references: UNIPROT:O39928; GB:Y13184
A;Experimental source: genotype 5a, which predominates in South Africa
A;Note: the translation of the nucleotide sequence is not complete in this paper
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin
F;2-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <EPM>
F;192-389/Product: major envelope protein E #status predicted <MEE>
F;384-408/Region: hypervariable #status predicted
F;390-730/Product: nonstructural protein NS1 #status predicted <NS1>
F;731-1007/Product: nonstructural protein NS2 #status predicted <NS2>
F;1008-1616/Product: hepatitis C virus genome polyprotein NS3 #status predicted <NS3>
F;1231-1238/Region: nucleotide-binding motif A (P-loop)
F;1313-1318/Region: nucleotide-binding motif B
F;1317-1320/Region: DEXH motif
F;1617-1863/Product: nonstructural protein NS4a #status predicted <N4A>
F;1864-2014/Product: nonstructural protein NS4b #status predicted <N4B>
F;2015-3014/Product: nonstructural protein NS5 #status predicted <NS5>
F;2210-2249/Region: interferon sensitivity determining #status predicted
Query Match 79.3%; Score 2511; DB 1; Length 3014;
Best Local Similarity 78.8%; Pred. No. 2.9e-184;
Matches 466; Conservative 53; Mismatches 72; Indels 0; Gaps 0;

```

```

QY 5 SMSYTWGALITPCAAEESKLPINALNSLRLHNLVYVSTTSRSASLRQKKVTFDRLOVL 64
DB 2424 SMSYTWGALITPCSAEEELKLPINPLSNYLLRHNLYVSTSSRAGLRQKKVTFDRLOVL 2483

QY 65 DDHYRDVLKEMKAKASTVYKAKLLSVEACKLTPPHSAKSKFGYGAOVRSLSRAVNHIR 124
DB 2484 DDHYREVVDKMRASKVKARLLPLEEACGLTPPHSARSKYGYGAKEVRSLDKALKHIE 2543

QY 125 SVWKDLLEDDTPTIOTTIMAKNEVFCVQPEKGGKRPARLIVFPDLGVRVCEKMALYDVVS 184
DB 2544 GVMODLLSDSDTPTIOTTIMAKNEVFCVQPEKGGKRPARLIVFPDLGVRVCEKMALYDVVAQ 2603

QY 185 TLPQAVMGSSYGFQYSPQORVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
DB 2604 KLPLAVMGSSYGFQYSPQORVDFLLKAKSKKIIMAFSYDTRCFDSTITEHDIITEESIY 2663

QY 245 QCCDLAPARQAIRSLTERLYVGGPMTNKSGQNGCYRRCRASGVLTTCGNTLTCYLKAA 304
DB 2664 QCCDLQPEARVAIRSLTQRLYCGGPMYNSKQCGYRRCRASGVLTTCGNTMTICYIKAL 2723

QY 305 AACRAAKLQDCTMLVNGDDLVIICESAGTQDDAASLRVFTAMTRYSAAPPDPPQPEYDL 364
DB 2724 ASCRAAKLRDCTLLVCGDDLVAICESQGTDEDEASLRFTAMTRYSAAPPDPPVPAYDL 2783

QY 365 ELITSCSSNVSAHDGAKRVYVLTDRPTVPLARAAMETARHTPVNSWLGNIIMVAPTLW 424
DB 2784 ELVTSCSSNVSAHDGAKRVYVLTDRPTVPLARAAMETARHTPVNSWLGNIIMVAPTLW 2843

QY 425 ARMLTMTHFFSILLAQEOLEKALDCQIYGACYSIEPLDLPQIIRLHGLSAFSLHSYSPG 484
DB 2844 ARIVLTMTHFFSVLQSQBOLEKTLAFEMVGVSVYVTPDLQPAIIQRLHGLSAFSLHSYSPS 2903

QY 485 EINRVASCLRLKGLVPPPLRVWHRARSVRAKLLSGGRAAICGKYLFWAVRTKLTPIIP 544
DB 2904 EINRVASCLRLKGLVPPPLRAWRHRAVRARAKLIAQGGRAAICGKYLFWAVRTKLTPLA 2963

QY 545 AASRLDLSGFWAGYSGGDIYHSLSRARPRWFMCLCLLLSVGVGIYLLPNR 595
DB 2964 DADRDLSSNFTVGAGGDIYHSHSRARPRNLLCULLLSVGVGIYLLPAR 3014

RESULT 10
GNMWJ8
genome polyprotein - hepatitis C virus (strain HC-J8)
N;Contains: capsid protein C; envelope protein M; hepatitis C virus genome polyprotein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C;Accession: A40250; PQ0397; PQ0559
R;Okamoto, H.; Kurai, K.; Okada, S.I.; Yamamoto, K.; Lizuka, H.; Tanaka, T.; Fukuda, S.;
Virology 188, 331-341, 1992
A;Title: Full-length sequence of a hepatitis C virus genome having poor homology to repo:
A;Reference number: A40250; MUID:92230232; PMID:1314459
A;Accession: A40250
A;Molecule type: genomic RNA
A;Residues: 1-3033 <OKA>
A;Cross-references: UNIPROT:P26661; GB:D10988; GB:D01221; NID:g221608; PIDN:BAA01761.1; I
J. Gen. Virol. 73, 1131-1141, 1992
A;Experimental source: isolate E-b12
R;Kato, N.; Ootsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimotohno,
Biochem. Biophys. Res. Commun. 181, 279-285, 1991
A;Title: Distribution of plural HCV types in Japan.
A;Reference number: PQ0393; MUID:92268871; PMID:1316939
A;Accession: PQ0397
A;Molecule type: genomic RNA
A;Residues: 2678-2754 <CHA>
A;Cross-references: DDBJ:D10134
A;Experimental source: isolate E-b12
R;Kato, N.; Ootsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimotohno,
Biochem. Biophys. Res. Commun. 181, 279-285, 1991
A;Title: Distribution of plural HCV types in Japan.
A;Reference number: PQ0554; MUID:92068204; PMID:1720309
A;Accession: PQ0559
A;Molecule type: mRNA
A;Residues: 2678-2729 <KAT>

```


A:Reference number: JQ1303; MUID:92044440; PMID:1658196

A:Accession: JQ1303

A:Molecule type: genomic RNA

A:Residues: 1-3033 <K>A>

A:Cross-references: UNIPROT:P26660; GB:D00944; NID:g221650; PIDN:BAA00792.1; PID:g221651

A:Experimental source: isolate HC-J6 from a Japanese individual

C:Superfamily: hepatitis C virus genome polyprotein

C:Keywords: ATP; glycoprotein; hydrolase; P-loop; polyprotein; serine proteinase; transmembrane

F:2-115/Product: capsid protein C #status predicted <CP>

F:116-191/Product: envelope protein M #status predicted <EP>

F:192-389/Product: major envelope protein E #status predicted <ME>

F:390-733/Product: nonstructural protein NS1 #status predicted <NS1>

F:734-1010/Product: nonstructural protein NS2 #status predicted <NS2>

F:1011-1619/Product: nonstructural protein NS3 #status predicted <NS3>

F:1619-1619/Product: hepatitis C virus NS3 #status predicted <NS3>

F:1316-1321/Region: nucleotide-binding motif B

F:1320-1323/Region: DEXH motif

F:1620-1866/Product: nonstructural protein NS4a #status predicted <NS4>

F:1867-2017/Product: nonstructural protein NS4b #status predicted <NS4>

F:2018-3033/Product: nonstructural protein NS5 #status predicted <NS>

F:196,209,234,305,325,417,423,430,448,477,534,542,538,578,627,649,1091,1217,1259,2038,281

Query Match	75.1%	Score 2380	DB 1	Length 3033
Best Local Similarity	75.6%	Pred. No. 3.5e-174		
Matches 447	Conservative 57	Mismatches 87	Indels 0	Gaps 0

Qy	5	SMSTVTGALITPCAAEESKLPINALSNSLRHNLVYTSRSASLSKQKVTFFORLQVL 64
Db	2443	SMSTVTGALITPCSPPEEKLPINPLNSLRLYHNKVCTTSKSASLRKAKVTFFORMQAL 2502
Qy	65	DDHYRDVLKEMKAKASTVTKALLSVEERACKLTPPHSAKSKGYGAKDVRSLSSRAVNHIR 124
Db	2503	DAHYSVLKDILKAASKVTARLLLTLEACQITPPHSAKSKGYGAKDVRSLSSRAVNHIR 2562
Qy	125	SVKDKLLEDTDPTQTTIMAKNEVFCVQPEKGGKRPARLIYVPPDLGVRVCEKMAALYDVVS 184
Db	2563	SVKDKLLEDTQPTPTTMAKNEVFCVDPTKGGKAARLIYVPPDLGVRVCEKMAALYDITQ 2622
Qy	185	TLPOAVMGSSYGFQYSPQRQVPEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
Db	2623	KLPOAVMGASYGFQYSPAQRVEFLKAWAEKKDPMGFSYDTRCFDSTVTENDIRVEESIY 2682
Qy	245	QCQDLAPARQATBSLTERLYVGGPMTNSKQNGYRCRASGLVTTSCGNTLTTCYLKAA 304
Db	2683	RACSLPEAHTAHSLSLTERLYVGGPMTNSKQNGYRCRASGLVTTSCGNTLTTCYLKAA 2742
Qy	305	AACRAAKLQDCTMLVNGDDLVIICESAGTQSDAASLRVFTTEAMTRYSPAGPDPPOPEYDL 364
Db	2743	AACKAGIIAPTLVCGDDLVIISQGTEDERNLRAFTTEAMTRYSPAGPDPPOPEYDL 2802
Qy	365	ELITSCSNVSAHDASGKRVYILTRDPTVPLARAAMETARHTPVNSLGNIMYAPTILW 424
Db	2803	ELITSCSNVSAVALGPQRRRYILTRDPTPIARAAMETVRHSVNSLGNIIQYAPTILW 2862
Qy	425	ARMILMTHFSSILLAQSOLEKALDCQYAGCYSTEPLDLPQIERLHGLSARLSHSYSG 484
Db	2863	ARMVLMTHFSSILMAQDLDQNLNFMFYGAVYSPLDLPALIERLHGLDFAFLSHLYTTPH 2922
Qy	485	EINRVASCLRLGVPPRLVVRHRRARSVRKLLSOGGRAAICGKYLFPNAVTKLKLTPP 544
Db	2923	ELTRVASALRKLGAPPRAMKSRARAVRASLISGGRAAVCGRYLFPNAVTKLKLTPP 2982
Qy	545	AASRLDLSGWFWAGYSGGDIYHSLSRARPWRFMCLCLLLSVGVGIYLLPNR 595
Db	2983	EARLLDLSSNFTVAGGDDIYHVSRRARPRLLLLGLLLFVGVGLFLLPAR 3033

RESULT 12

JQ0879

NS5 protein - hepatitis C virus (strain J4) (fragment)

C:Species: hepatitis C virus

C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jul-2004

C:Accession: JQ0879

R:Okamoto, H.

submitted to JIPID, January 1991
A:Description: The 5'-terminal and 3'-terminal sequences of the genomic RNA of hepatitis
A:Reference number: JQ0879
A:Accession: JQ0879
A:Molecule type: genomic RNA
A:Residues: 1-365 <OK>
A:Cross-references: UNIPROT:Q81717
A:Experimental source: strain J4
C:Superfamily: hepatitis C virus genome polyprotein

Query Match 59.6%; Score 1886; DB 2; Length 365;
Best Local Similarity 96.7%; Pred. No. 2.1e-137;
Matches 353; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

QY 231 TVTENDIRVESIYQCCDLAPARQAIRSLTERLYVGGPMTNSKGQNGYRCRASGVLT 290
DB 1 TVTESDIRVESIYQCCDLAPARQAIRSLTERLYVGGPMTNSKGQNGYRCRASGVLT 60

QY 291 TSCGNLTTCYLKAAACRAAKLQDCTMLVNGDDLVIICESAGTQEDDAASLRVFTTEAMTRY 350
DB 61 TSCGNLTTCYLKATACRAAKLQDCTMLVNGDDLVIICESAGTQEDDAASLRVFTTEAMTRY 120

QY 351 SAPGDPDPQPEYDLELITSCSNVSAHDASGKRVYLLTRDPTVPLARAAWETARHTPVN 410
DB 121 SAPGDPDPQPEYDLELITSCSNVSAHDASGKRVYLLTRDPTVPLARAAWETARHTPVN 180

QY 411 SWLGNIMYAPTLWARMLTMTHFFSILLAOQLKALDCQIYGACYSIEPLDLQIIRL 470
DB 181 SWLGNIMYAPALWARMLTMTHFFSILLAOQLKALDCQIYGACYSIEPLDLQIIRL 240

QY 471 HGLSAFSLHSYSGEINRVASCLRKLGVPPLRVHRARSVRAKLLSOGGAAACGKYL 530
DB 241 HGLSAFSLHSYSGEINRVASCLRKLGVPPLRVHRARSVRAKLLSOGGAAACGKYL 300

QY 531 NWAIVTKLKLTPIPAASRLDLSGMFVAGYSGGDIYHLSLRARPRFWMCLLLLSVGVGIY 590
DB 301 NWAIVTKLKLTPIPAASRLDLSGMFVAGYSGGDIYHLSLRARPRFWMCLLLLSVGVGIY 360

QY 591 LLPNR 595
DB 361 LLPNR 365

RESULT 13
JQ0880
NS5 protein - hepatitis C virus (strain J1) (fragment)
C:Species: hepatitis C virus
C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jul-2004
C:Accession: JQ0880
R:Okamoto, H.
submitted to JIPID, January 1991
A:Description: The 5'-terminal and 3'-terminal sequences of the genomic RNA of hepatitis
A:Reference number: JQ0879
A:Accession: JQ0880
A:Molecule type: genomic RNA
A:Residues: 1-365 <OK>
A:Cross-references: UNIPROT:Q81716
A:Experimental source: strain J1
C:Superfamily: hepatitis C virus genome polyprotein

Query Match 55.3%; Score 1750; DB 2; Length 365;
Best Local Similarity 88.2%; Pred. No. 6.1e-127;
Matches 322; Conservative 23; Mismatches 20; Indels 0; Gaps 0;

QY 231 TVTENDIRVESIYQCCDLAPARQAIRSLTERLYVGGPMTNSKGQNGYRCRASGVLT 290
DB 1 TVTESDIRTEAIYQCCDLDPQARVAIKSLTERLYVGGPMTNSRGNGCYRCRASGVLT 60

QY 291 TSCGNLTTCYLKAAACRAAKLQDCTMLVNGDDLVIICESAGTQEDDAASLRVFTTEAMTRY 350
DB 61 TSCGNLTTCYLKAAACRAAKLQDCTMLVNGDDLVIICESAGTQEDDAASLRVFTTEAMTRY 120

QY 351 SAPGDPDPQPEYDLELITSCSNVSAHDASGKRVYLLTRDPTVPLARAAWETARHTPVN 410

DB 121 SAPGDPDPQPEYDLELITSCSNVSAHDGKRVYLLTRDPTVPLARAAWETARHTPVN 180

QY 411 SWLGNIMYAPTLWARMLTMTHFFSILLAOQLKALDCQIYGACYSIEPLDLQIIRL 470
DB 181 SWLGNIMYAPTLWARMLTMTHFFSILLAOQLKALDCQIYGACYSIEPLDLQIIRL 240

QY 471 HGLSAFSLHSYSGEINRVASCLRKLGVPPLRVHRARSVRAKLLSOGGAAACGKYL 530
DB 241 HGLSAFSLHSYSGEINRVASCLRKLGVPPLRVHRARSVRAKLLSOGGAAACGKYL 300

QY 531 NWAIVTKLKLTPIPAASRLDLSGMFVAGYSGGDIYHLSLRARPRFWMCLLLLSVGVGIY 590
DB 301 NWAIVTKLKLTPIPAASRLDLSGMFVAGYSGGDIYHLSLRARPRFWMCLLLLSVGVGIY 360

QY 591 LLPNR 595
DB 361 LLPNR 365

RESULT 14
JQ0883
Genome polyprotein - hepatitis C virus (strain J7) (fragments)
N:Contains: NS5 protein
C:Species: hepatitis C virus
C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jul-2004
C:Accession: JQ0883
R:Okamoto, H.
submitted to JIPID, January 1991
A:Description: The 5'-terminal and 3'-terminal sequences of the genomic RNA of hepatitis
A:Reference number: JQ0879
A:Accession: JQ0883
A:Molecule type: genomic RNA
A:Residues: 1-874 <OK>
A:Cross-references: UNIPROT:Q7LZY6
A:Experimental source: strain J7
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: polyprotein
C:Product: NS5 protein (fragment) #status predicted <NS5>
F:510-874/

Query Match 46.5%; Score 1471.5; DB 2; Length 874;
Best Local Similarity 70.0%; Pred. No. 5.6e-105;
Matches 282; Conservative 40; Mismatches 64; Indels 17; Gaps 3;

QY 195 YGFOYSPKQVFEFLVNTWAKKCPMGF-SYDTRCFD-STVTENDIRVEESIYQCCDLAPE 252
DB 487 YCWHYPPK-----PCGIVSAKTVCGPVYTVTERDIRTEESIYQACSLPQE 531

QY 253 AROAIRSLTERLYVGGPMTNSKGQNGYRCRASGVLTTCGNTLTTCYLKAAACRAAKL 312
DB 532 ARTAIHSLTERLYVGGPMTNSKGQNGYRCRASGVLTTSIGNTWTCYIKALAAKANGI 591

QY 313 QDCTMLVNGDDLVIICESAGTQEDDAASLRVFTTEAMTRYAPPDPQPEYDLELITSCS 372
DB 592 KDPIMLVCGDDLVIISQNGNEEDERNLRAFTTEAMTRYAPPDLPPEYDLELITSCS 651

QY 373 NVSAHDASGKRVYLLTRDPTVPLARAAWETARHTPVNSWLNIGMIIMYAPTILWAKMLMTH 432
DB 652 NVSVALDPGRGRYLLTRDPTTISRAAWETVRHSPVNSWLNIGMIIMYAPTILWAKMLMTH 711

QY 433 FFSILLAOQLKALDCQIYGACYSIEPLDLQIIRLHGLSAFSLHSYSGEINRVAS 492
DB 712 FFNILLAOQTLNQNLNEMFYGAVSVNPLDLPALIERLHGLDAFSLHTYSPEHSRVNAT 771

QY 493 LRKLGVPLPLRVHRARSVRAKLLSOGGAAACGKYLFFNWAIVTKLKLTPIPAASRLDLS 552
DB 772 LRKLGAFFLRKRSRARSRAVRAASLIAQGGAAACGKYLFFNWAIVTKLKLTPLEAARDLS 831

QY 553 GWFVAGYSGGDIYHLSLRARPRFWMCLLLLSVGVGIYLLPNR 595
DB 832 GWFTVAGGGGDIFHSVSHARPRLLLLCLLLSVGVGIFLLPAR 874

Search completed: September 22, 2005, 14:48:09
Job time : 22 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2005, 14:36:26 ; Search time 78 Seconds
(without alignments)
3945.636 Million cell updates/sec

Title: US-10-712-479-2

Perfect score: 3167

Sequence: 1 MASMSMTWTGALITPCAA.....LLSVGVGYLLPNRRHHHH 601

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3040	96.0	3010	2 Q9DTE6	Q9dte6 hepatitis c
2	3040	96.0	3010	2 Q9J310	Q9j310 hepatitis c
3	3038	96.0	3010	2 Q93077	Q93077 hepatitis c
4	3038	95.9	3010	2 Q9DTE9	Q9dte9 hepatitis c
5	3034	95.8	3010	2 Q81760	Q81760 hepatitis c
6	3030	95.7	3010	2 Q02829	Q02829 hepatitis c
7	3027	95.6	3010	2 Q68285	Q68285 hepatitis c
8	3027	95.6	3010	2 Q81825	Q81825 hepatitis c
9	3026	95.5	3010	2 Q9J3H0	Q9j3h0 hepatitis c
10	3025	95.5	3010	2 Q9QIX5	Q9qix5 hepatitis c
11	3024	95.5	3010	2 Q9QIX5	Q9qix5 hepatitis c
12	3022	95.4	3010	2 P89966	P89966 hepatitis c
13	3022	95.4	3014	2 Q09796	Q09796 hepatitis c
14	3021	95.4	3010	2 Q09796	Q09796 hepatitis c
15	3020	95.4	591	2 Q39930	Q39930 hepatitis c
16	3020	95.4	591	2 Q80M22	Q80m22 hepatitis c
17	3019	95.3	591	2 Q80M22	Q80m22 hepatitis c
18	3018	95.3	591	2 Q9JGN9	Q9jgn9 hepatitis c
19	3018	95.3	3010	2 Q68826	Q68826 hepatitis c
20	3018	95.3	3010	2 Q807P3	Q807p3 hepatitis c
21	3017	95.3	591	2 Q80M70	Q80m70 hepatitis c
22	3016	95.2	2284	2 Q81817	Q81817 hepatitis c
23	3016	95.2	3010	1 POLG_HCVBK	P26663 h genome po
24	3016	95.2	3010	1 POLG_HCVJA	P26663 h genome po
25	3016	95.2	3010	2 Q9DTE4	Q9dte4 hepatitis c
26	3015	95.2	591	2 Q80MH3	Q80mh3 hepatitis c
27	3015	95.2	591	2 Q80M10	Q80m10 hepatitis c
28	3015	95.2	3010	2 Q82969	Q82969 hepatitis c
29	3015	95.2	3010	2 Q92971	Q92971 hepatitis c
30	3015	95.2	3010	2 Q92972	Q92972 hepatitis c
31	3015	95.2	3010	2 Q9J3H9	Q9j3h9 hepatitis c

RESULT 1

ID	Q9DTE6	PRELIMINARY;	PRT;	3010 AA.	
AC	Q9DTE6;				Q99au2 hepatitis c
DT	01-MAR-2001 (Tremblrel. 16, Created)				Q9j3f9 hepatitis c
DT	01-MAR-2001 (Tremblrel. 16, Last sequence update)				Q02828 hepatitis c
DT	01-MAR-2004 (Tremblrel. 26, Last annotation update)				Q9dcd9 hepatitis c
DE	Polyprotein.				P90194 hepatitis c
OS	Hepatitis C virus.				Q9j3h7 hepatitis c
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;				P90193 hepatitis c
OC	Hepacivirus.				Q80mh2 hepatitis c
OX	NCBI_TaxID=11103;				Q80mh5 hepatitis c
RN	[1]				Q92970 hepatitis c
RP	SEQUENCE FROM N.A.				Q00269 h genome po
RC	TISSUE=Serum;				Q9qiy3 hepatitis c
RA	Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,				Q9qiy4 hepatitis c
RA	Hatahara T., Ohka Y., Kanai K., Maruo H., Baba K., Hijikata M.,				
RA	Mishiro S.;				
RT	"Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients				
RT	with hepatocellular carcinoma: the 'progression score' revisited.";				
RL	Hepatol. Res. 20:161-171(2001).				
DR	EMBL; AB049091; BAB18804.1; -.				
DR	PIR; A61196; A61196.				
DR	PIR; PS0329; PS0329.				
DR	HSSP; Q81755; 1DXP.				
DR	GO; GO:0016021; C:integral to membrane; IEA.				
DR	GO; GO:0019028; C:viral capsid; IEA.				
DR	GO; GO:0019031; C:viral envelope; IEA.				
DR	GO; GO:0005534; F:ATP binding; IEA.				
DR	GO; GO:0008036; F:ATP-dependent helicase activity; IEA.				
DR	GO; GO:0016787; F:hydrolase activity; IEA.				
DR	GO; GO:0003723; F:RNA binding; IEA.				
DR	GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.				
DR	GO; GO:0008236; F:serine-type peptidase activity; IEA.				
DR	GO; GO:0005198; F:structural molecule activity; IEA.				
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.				
DR	GO; GO:0006350; P:transcription; IEA.				
DR	GO; GO:0019079; P:viral genome replication; IEA.				
DR	GO; GO:0019087; P:viral transformation; IEA.				
DR	InterPro; IPR000345; CytC_heme_BS.				
DR	InterPro; IPR001410; DEAD.				
DR	InterPro; IPR011545; DEAD/DEAH N.				
DR	InterPro; IPR002522; HCV capsid.				
DR	InterPro; IPR002519; HCV_core.				
DR	InterPro; IPR002531; HCV_NSI.				
DR	InterPro; IPR000745; HCV_NS4a.				
DR	InterPro; IPR001490; HCV_NS4b.				
DR	InterPro; IPR002868; HCV_NS5a.				
DR	InterPro; IPR002166; HCV_RdRP.				
DR	InterPro; IPR001650; Helicase_C.				
DR	InterPro; IPR004109; Peptidase_S29.				
DR	InterPro; IPR009003; Pept_Ser_Cys.				
DR	InterPro; IPR002518; Pept_U39_HCV_NS2.				

ALIGNMENTS

DR	InterPro; IPR007095; RNA pol DS PS.	OS	Hepatitis C virus.
DR	InterPro; IPR007094; RNA_pol_PSVir.	OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
DR	Pfam; PF01543; HCV_capsid; 1.	OC	Hepacivirus.
DR	Pfam; PF01539; HCV_env; 1.	OX	NCBI_TaxID=11103;
DR	Pfam; PF01560; HCV_NS1; 1.	RN	[1]
DR	Pfam; PF01538; HCV_NS2; 1.	RP	SEQUENCE FROM N.A.
DR	Pfam; PF02907; HCV_NS3; 1.	RC	STRAIN=MD12;
DR	Pfam; PF01006; HCV_NS4a; 1.	RA	Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
DR	Pfam; PF01001; HCV_NS4b; 1.	RL	Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR	Pfam; PF01506; HCV_NS5a; 1.	DR	EMBL; AF207753; AAF65943.1; -.
DR	Pfam; PF00271; Helicase C; 1.	DR	PIR; A61196; A61196.
DR	Pfam; PF00998; Viral RdRP; 1.	DR	PIR; PQ0246; PQ0246.
DR	SMART; SM00487; DEXDC; 1.	DR	PIR; PS0329; PS0329.
DR	SMART; SM00487; HELIC; 1.	DR	HSP; O8JYS1; 1CW.
DR	PROSITE; PS00190; CYTOCHROME_C; UNKNOWN 1.	DR	GO; GO:0016021; C:integral to membrane; IEA.
KW	ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase;	DR	GO; GO:0019028; C:viral capsid; IEA.
KW	Hydrolase; Nonstructural protein; Transmembrane.	DR	GO; GO:0019031; C:viral envelope; IEA.
SQ	SEQUENCE 3010 AA; 327043 MW; 3807DC6879684C95 CRC64;	DR	GO; GO:0005524; F:ATP binding; IEA.
		DR	GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
		DR	GO; GO:0003723; F:RNA binding; IEA.
		DR	GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
		DR	GO; GO:0008236; F:serine-type peptidase activity; IEA.
		DR	GO; GO:0005198; F:structural molecule activity; IEA.
		DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
		DR	GO; GO:0006350; P:transcription; IEA.
		DR	GO; GO:0019079; P:viral genome replication; IEA.
		DR	GO; GO:0019087; P:viral transformation; IEA.
		DR	InterPro; IPR000345; CytC_heme_BS.
		DR	InterPro; IPR001410; DEAD.
		DR	InterPro; IPR011545; DEAD/DEAH N.
		DR	InterPro; IPR002522; HCV capsid.
		DR	InterPro; IPR002521; HCV_core.
		DR	InterPro; IPR002519; HCV_env.
		DR	InterPro; IPR002531; HCV_NS1.
		DR	InterPro; IPR000745; HCV_NS4a.
		DR	InterPro; IPR001490; HCV_NS4b.
		DR	InterPro; IPR002868; HCV_NS5a.
		DR	InterPro; IPR002166; HCV_RdRP.
		DR	InterPro; IPR001650; Helicase_C.
		DR	InterPro; IPR004109; Peptidase_S29.
		DR	InterPro; IPR009003; Pept_Ser_Cys.
		DR	InterPro; IPR002518; Pept_U39_HCV_NS2.
		DR	InterPro; IPR007095; RNA_pol_DS_PS.
		DR	InterPro; IPR007094; RNA_pol_PSVir.
		DR	Pfam; PF01543; HCV_capsid; 1.
		DR	Pfam; PF01542; HCV_core; 1.
		DR	Pfam; PF01539; HCV_env; 1.
		DR	Pfam; PF01560; HCV_NS1; 1.
		DR	Pfam; PF01538; HCV_NS2; 1.
		DR	Pfam; PF02907; HCV_NS3; 1.
		DR	Pfam; PF01006; HCV_NS4a; 1.
		DR	Pfam; PF01001; HCV_NS4b; 1.
		DR	Pfam; PF01506; HCV_NS5a; 1.
		DR	Pfam; PF00271; Helicase_C; 1.
		DR	Pfam; PF00998; Viral RdRP; 1.
		DR	SMART; SM00487; DEXDC; 1.
		DR	PROSITE; PS00190; CYTOCHROME_C; UNKNOWN 1.
		DR	Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW	Polyprotein; Transmembrane.	KW	Polyprotein; Transmembrane.
SQ	SEQUENCE 3010 AA; 326693 MW; 074098DB305AF1A9 CRC64;	SQ	SEQUENCE 3010 AA; 326693 MW; 074098DB305AF1A9 CRC64;
			Query Match 96.0%; Score 3040; DB 2; Length 3010;
			Best Local Similarity 97.3%; Pred. No. 6.3e-228;
			Matches 574; Conservative 11; Mismatches 6; Indels 0; Gaps 0;
Qy	5 SMSYTWTCALITPCAABESKLPINALNSLRHNLVYSTTSRSASLRQKQKVTDFRLQVL 64	Qy	5 SMSYTWTCALITPCAABESKLPINALNSLRHNLVYSTTSRSASLRQKQKVTDFRLQVL 64
Db	2420 SMSYTWTCALITPCAABESKLPINALNSLRHNLVYSTTSRSASLRQKQKVTDFRLQVL 2479	Db	2420 SMSYTWTCALITPCAABESKLPINALNSLRHNLVYSTTSRSASLRQKQKVTDFRLQVL 2479
Qy	65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAADVRSLSRAVNHR 124	Qy	65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAADVRSLSRAVNHR 124
Db	2480 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAADVRSLSRAVNHR 2539	Db	2480 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAADVRSLSRAVNHR 2539
Qy	125 SVWKDLLEDTPTQTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCCKMALYDVVS 184	Qy	125 SVWKDLLEDTPTQTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCCKMALYDVVS 184
Db	2540 SVWKDLLEDTPTQTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCCKMALYDVVS 2599	Db	2540 SVWKDLLEDTPTQTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCCKMALYDVVS 2599
Qy	185 TLPQAVMGSSYGFQSPQGRVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVESIY 244	Qy	185 TLPQAVMGSSYGFQSPQGRVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVESIY 244
Db	2600 TLPQAVMGSSYGFQSPQGRVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVESIY 2659	Db	2600 TLPQAVMGSSYGFQSPQGRVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVESIY 2659
Qy	245 QCCDLAPEARQAIRSLTERLYVGGPMNTNSKQNGCYRRCRASGVLTTCGNTLTCYLKAA 304	Qy	245 QCCDLAPEARQAIRSLTERLYVGGPMNTNSKQNGCYRRCRASGVLTTCGNTLTCYLKAA 304
Db	2660 QCCDLAPEARQAIRSLTERLYVGGPMNTNSKQNGCYRRCRASGVLTTCGNTLTCYLKAT 2719	Db	2660 QCCDLAPEARQAIRSLTERLYVGGPMNTNSKQNGCYRRCRASGVLTTCGNTLTCYLKAT 2719
Qy	305 AACRAAKLQDCTMLVNGDDLVIICESAGTQDAASLRVFTAMTRYASAPGDPQPEYDL 364	Qy	305 AACRAAKLQDCTMLVNGDDLVIICESAGTQDAASLRVFTAMTRYASAPGDPQPEYDL 364
Db	2720 AACRAAKLQDCTMLVNGDDLVIICESAGTQDAASLRVFTAMTRYASAPGDPQPEYDL 2779	Db	2720 AACRAAKLQDCTMLVNGDDLVIICESAGTQDAASLRVFTAMTRYASAPGDPQPEYDL 2779
Qy	365 ELITSCSNVSVAHDASGRVYVYLTTRDPTVPLAARAWEETARHTPVNSWLGNIIMVAPTLW 424	Qy	365 ELITSCSNVSVAHDASGRVYVYLTTRDPTVPLAARAWEETARHTPVNSWLGNIIMVAPTLW 424
Db	2780 ELITSCSNVSVAHDASGRVYVYLTTRDPTVPLAARAWEETARHTPVNSWLGNIIMVAPTLW 2839	Db	2780 ELITSCSNVSVAHDASGRVYVYLTTRDPTVPLAARAWEETARHTPVNSWLGNIIMVAPTLW 2839
Qy	425 ARMLIMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484	Qy	425 ARMLIMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
Db	2840 ARMLIMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2899	Db	2840 ARMLIMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2899
Qy	485 EINRVASCLRLKGVPLVWHRHARSVRKALLSQGRAAICCKGLFNWAVRTKLTPIP 544	Qy	485 EINRVASCLRLKGVPLVWHRHARSVRKALLSQGRAAICCKGLFNWAVRTKLTPIP 544
Db	2900 EINRVASCLRLKGVPLVWHRHARSVRKALLSQGRAAICCKGLFNWAVRTKLTPIP 2959	Db	2900 EINRVASCLRLKGVPLVWHRHARSVRKALLSQGRAAICCKGLFNWAVRTKLTPIP 2959
Qy	545 AASRLDLGSGFWAGVSGGDIYHLSRARPRFVLCILLLSVGVGIYLLPNR 595	Qy	545 AASRLDLGSGFWAGVSGGDIYHLSRARPRFVLCILLLSVGVGIYLLPNR 595
Db	2960 AASRLDLGSGFWAGVSGGDIYHLSRARPRFVLCILLLSVGVGIYLLPNR 3010	Db	2960 AASRLDLGSGFWAGVSGGDIYHLSRARPRFVLCILLLSVGVGIYLLPNR 3010
RESULT 2			
Q9J310	PRELIMINARY; PRT; 3010 AA.		
ID	Q9J310		
AC	Q9J310;		
DT	01-OCT-2000 (TremBLrel. 15, Created)		
DT	01-OCT-2000 (TremBLrel. 15, Last sequence update)		
DT	01-MAR-2004 (TremBLrel. 26, Last annotation update)		
DE	Polyprotein.		

```
QY 125 SVWKDLLEDDTPTIOTTIMAKNEVFCVQPEKGGKRPALIVFPDGLGVRCCKMALYDVVS 184
Db 2540 SVWKDLLEDDTPTIOTTIMAKNEVFCVQPEKGGKRPALIVFPDGLGVRCCKMALYDVVS 2599
QY 185 TLPQAVMGSSYGFQYSPKQRFVFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
Db 2600 TLPQAVMGSSYGFQYSPKQRFVFLVNTWKSKCPMGFSYDTRCFDSTVTENDIRVEESIY 2659
QY 245 QCCDLAPPEARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTSCGNTLTCLYKAA 304
Db 2660 QCCDLAPPEARQIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTSCGNTLTCLYKAS 2719
QY 305 AACRAAKLQDCTMLVNGDDLVIICESAGTQDAASLRVFTTEAMTRYSPAPGDPPOPEYDL 364
Db 2720 AACRAAKLQDCTMLVNGDDLVIICESAGTQDAASLRVFTTEAMTRYSPAPGDPPOPEYDL 2779
QY 365 ELITSCSSNSVVAHDASGKRVYILTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTLM 424
Db 2780 ELITSCSSNSVVAHDASGKRVYILTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTLM 2839
QY 425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAPSLHSYSPG 484
Db 2840 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAPSLHSYSPG 2899
QY 485 E1NRVASCLRLKGLVPPPLRVWRHRSVRAKLLSQGGRAAICGKYLFWNAVTKLKLTPIP 544
Db 2900 E1NRVASCLRLKGLVPPPLRVWRHRSVRAKLLSQGGRAATCGKYLFWNAVTKLKLTPIP 2959
QY 545 AASRLDLSGFWVAGYSGGDIYHLSLRARPRFMLCLLLSVGVGYLLPNR 595
Db 2960 AASRLDLSGFWVAGYSGGDIYHLSLRARPRFMLCLLLSVGVGYLLPNR 3010

RESULT 3
Q93077 PRELIMINARY; PRT; 3010 AA.
AC Q93077;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98122498; PubMed=9462666; DOI=10.1002/hep.510270242;
RA Aizaki H., Aoki Y., Harada T., Ishii K., Suzuki T., Nagamori S.,
RA Toda G., Matsuura Y., Miyamura T.;
RT "Full-length complementary DNA of hepatitis C virus genome from an
RT infectious blood sample.";
RL Hepatology 27:621-627(1998).
DR EMBL; D89815; BAA25076.1; -.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSSP; Q8JYS1; ICWX.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
```

```
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39 HCV NS2.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol PSvir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; Viral RdRP; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3010 AA; 327352 MW; 888BBA102A733390 CRC64;
```

```
Query Match 96.0%; Score 3039; DB 2; Length 3010;
Best Local Similarity 97.6%; Pred. No. 7.5e-228;
Matches 577; Conservative 7; Mismatches 7; Indels 0; Gaps 0;
QY 5 SMSYTWGALITPCAABESKLPINALSNLSLRHNLVYSTTSRSASLRQKKVTFDRQLVL 64
Db 2420 SMSYTWGALITPCAABESKLPINALSNPLLRHNMVYSTTSRSASLRQKKVTFDRMQLV 2479
QY 65 DDHYRDVLKEMKAKASTVAKLLSVEEACKLTPPHSAKSKFGYGAKVRSLSRAVNHIR 124
Db 2480 DDHYRDVLKEMKAKASTVAKLLSVEEACKLTPPHSAKSKFGYGAKVRSLSRAVNHIR 2539
QY 125 SVWKDLLEDDTPTIOTTIMAKNEVFCVQPEKGGKRPALIVFPDGLGVRCCKMALYDVVS 184
Db 2540 SVWKDLLEDDTPTIOTTIMAKNEVFCVQPEKGGKRPALIVFPDGLGVRCCKMALYDVVS 2599
QY 185 TLPQAVMGSSYGFQYSPKQRFVFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
Db 2600 TLPQAVMGSSYGFQYSPKQRFVFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 2659
QY 245 QCCDLAPPEARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTSCGNTLTCLYKAA 304
Db 2660 QCCDLAPPEARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTSCGNTLTCLYKAA 2719
QY 305 AACRAAKLQDCTMLVNGDDLVIICESAGTQDAASLRVFTTEAMTRYSPAPGDPPOPEYDL 364
Db 2720 AACRAAKLQDCTMLVNGDDLVIICESAGTQDAASLRVFTTEAMTRYSPAPGDPPOPEYDL 2779
QY 365 ELITSCSSNSVVAHDASGKRVYILTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTLM 424
Db 2780 ELITSCSSNSVVAHDASGKRVYILTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTLM 2839
QY 425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAPSLHSYSPG 484
Db 2840 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAPSLHSYSPG 2899
QY 485 E1NRVASCLRLKGLVPPPLRVWRHRSVRAKLLSQGGRAAICGKYLFWNAVTKLKLTPIP 544
Db 2900 E1NRVASCLRLKGLVPPPLRVWRHRSVRAKLLSQGGRAATCGKYLFWNAVTKLKLTPIP 2959
```

Qy	545 AASRLDLSGWFVAGYSGGDIIYHSLSRARPRWFMCLLLLLSVGVGYLLPNR 595 ---:--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
D6	2960 EASQLDLSGWFVAGYSGGDIIYHSLSRARPRWFMCLLLLLSVGVGYLLPNR 3010

RESULT 4

Q9DTE9	PRELIMINARY; PRT; 3010 AA.
AD	Q9DTE9
AC	Q9DTE9
DT	01-WAR-2001 (TrEMBLrel. 16, Created)
DT	01-WAR-2001 (TrEMBLrel. 16, Last sequence update)
DT	01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE	Polyprotein.
DE	Hepatitis C virus.
OS	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
OC	Hepacivirus.
NCBI_TaxID=111103;	
ON	[1]
RN	SEQUENCE FROM N.A.
RC	TISSUE=Serum;
RA	Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K., Hatanahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijioka M., Mishiro S.;
RA	"Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients with hepatocellular carcinoma: the 'progression score' revisited.";
RT	Hepacit. Res. 20:161-171(2001).
RL	EMBL; AB043088; BAB18801.1; -.
DR	PIR; A61196; A61196.
DR	PIR; PQ0246; PQ0246.
DR	PIR; PQ0804; PQ0804.
DR	PIR; PS0329; PS0329.
DR	HSP; Q8JYS1; 1CW.
DR	GO; GO:0016021; C:integral to membrane; IEA.
DR	GO; GO:0019028; C:viral capsid; IEA.
DR	GO; GO:0019031; C:viral envelope; IEA.
DR	GO; GO:0005524; F:ATP binding; IEA.
DR	GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR	GO; GO:0003723; F:RNA binding; IEA.
DR	GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR	GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR	GO; GO:0006350; F:proteolysis and peptidolysis; IEA.
DR	GO; GO:0006350; F:transcription; IEA.
DR	GO; GO:0019079; P:viral genome replication; IEA.
DR	GO; GO:0019087; P:viral transformation; IEA.
DR	InterPro; IPR000345; CytC_heme_BS.
DR	InterPro; IPR001410; DEAD.
DR	InterPro; IPR011545; DEAD/DEAH_N.
DR	InterPro; IPR002522; HCV capsid.
DR	InterPro; IPR002521; HCV core.
DR	InterPro; IPR002519; HCV env.
DR	InterPro; IPR002531; HCV NS1.
DR	InterPro; IPR000745; HCV NS4a.
DR	InterPro; IPR001490; HCV NS4b.
DR	InterPro; IPR002868; HCV NS5a.
DR	InterPro; IPR002166; HCV RdRp.
DR	InterPro; IPR001650; Helicase_C.
DR	InterPro; IPR004109; Peptidase_S29.
DR	InterPro; IPR009003; Pept_Ser_Cys.
DR	InterPro; IPR002518; Pept_U39_HCV_NS2.
DR	InterPro; IPR007095; RNA_pol_DS_P8.
DR	InterPro; IPR007094; RNA_pol_Fsvir.
DR	Pfam; PF01543; HCV capsid; 1.
DR	Pfam; PF01542; HCV core; 1.
DR	Pfam; PF01539; HCV env; 1.
DR	Pfam; PF01560; HCV NS1; 1.
DR	Pfam; PF01538; HCV NS2; 1.
DR	Pfam; PF02307; HCV NS3; 1.
DR	Pfam; PF01006; HCV NS4a; 1.
DR	Pfam; PF01001; HCV NS4b; 1.
DR	Pfam; PF01506; HCV NS5a; 1.
DR	Pfam; PF00271; Helicase_C; 1.
DR	Pfam; PF00998; Viral_RdRp; 1.

DR SMART; SM00487; DEXDc; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3010 AA; 326780 MW; 668CFFEA5FEC3658 CRC64;

Query Match 95.9%: Score 3038; DB 2; Length 3010;

Best Local Similarity 96.8%; Pred. NO. 9e-228;
Matches 572; Conservative 13; Mismatches 6; Indels 0; Caps 0;

QY	5	SMSYTWTCALITPCAAAEESKLPINALSNSLLRHNLVYSTTSSASLRQKKVTFDRLOVL	64
DB	2420	SMSYTWTCALITPCAAAEESKLPINALSNSLLRHNLMIYATTSRSAGLRQKKVTFDRLOVL	2479

Qy 65 DDHYDVLKEMKAKASTWKALLSVEBACLTPPHSAKSGFYGAKDVRLSSRAVNHIR 124

Db 2480 DDHYRDVLKEMRAKASTVKALLSVEEACKLTTPPHSAKSKFGYGAKDVRNLSSKAVNHIR 25339

Qy 125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGRKPARLIVFDLGVRCERKALYDVVS 184

DB 2540 SVWKOLLEDSERPIDTITMARNEIFCQVPEKGRKPARIIVFPDLGVRCVERMALYDWS 2599

C: 1 0 E TTY DOAMTCCEVCVOVSVDVGBUJDEI VNTMUKA KZCDMCSSVDHCBENCSMTNINIDUBESTV 244

[illegible]

245 QV OCDDLAPAFARQAIRSLTERLVYGGPPTNNSKQNGCYRRCRAGSVLTTCSCGNLTTCYV.KAA 304

Db 2660 QCCDLAPEARQAIRSLTERLYYGGPLTNSKQNGCYRRCASGVLTTSCGNTLTCLYKAS 2719

QY 305 AACRAAKIQDCTMLVNGDDLWVICESAGTQEDAAASLRFTEAMTRYSAAPGPPQPEYDL 364

Db 2720 AACRAAKLQDCTMLVNGDDLVI CESAGTQEDAASLRAFTEAMTRYSA PGPQPPEYDL 2779

QY 365 ELITSCSSNVSAHDAGKRVYLTTRDPTVPLARAAWETARTPTVNSWLGNTIMYAPTJW 424

Db 2780 ELITSCSSNVSAHDASCKRVYLLTRDPTTPLARAANETARHTPVNSWLGNIIMYAPTLW 2839

QY 425 ARMI L T W F F S I L L A Q E L K A L D C Q I Y G A C Y S I E P L D L P Q I E R L H G L S A F S L H S Y S P G 484

Db 2840 ARWILMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLQPIERLHGLSAFSLHSYSPG 2899

QY 485 E1NRVASCRLKLGVPPLRVVRHRSVRAKLLSQGGRAAICGKYLFWAVRTHKLTPIP 544

DB	2900	EINRVASCTKRLGVPPPLVRVRRHRSVSEAKLILSQGSRRAATCGKLYLFNVAVKTKRLTFLP	29359
Q:	E45	AAQDIDSCWEYACVSCCNIVHSI SDAPDDWEMI CTILISVQVCIIVLDND	595

QY ASAKKLIDJSGWFWAGISGGUIIISLSRKAPRWFMLCLLLLSVSGVGIILLFNN 333

QY

Dp ASRKLPIISGFWAGYSGSDIYHLSLRAPRWFMLCLLLLSVSGVGIYLLPNN 3010

[illegible]

RESULT 5
Q81760

ID	Q81760	PRELIMINARY;	PRT; 3010 AA.
AC	Q81760;		

DT	01-NOV-1996	(TREM)rel. 01, Created
DT	01-NOV-1996	(TREM)rel. 01, Last sequence update)
DT	01-MAR-2004	(TREM)rel. 02, Last annotation update)

DT 01-MAR-2004 (FEMBLrel. 26, Last annotation update)
DE Polypeptin.
OS *Haemophilus C virus*

OS Hepacivirus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OC Hepacivirus.

OX	NCBI_TaxID=11103;
RN	[1]

RP SEQUENCE FROM N.A.
RX MEDLINE=93359897; PubMed=8394876;

RA Wang Y., Okamoto H., Tsuda F., Nagayama R., Tao Q.M., Mishiro S.;
RT "Prevalence, genotypes, and an isolate (HC-C2) of hepatitis C virus in

RT Chinese patients with liver disease.";
RL J. Med. Virol. 40:254-260(1993).
DE EMBL. M10934. BAB01728 1. -

DK EMBL; D10934; BAF0120.1; -.
DR PIR; A61196; A61196.
DR PIR; P00246; P00246.

[illegible]


```

DR PIR; PQ0254; PQ0254.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSSP; Q8JYS1; 1CW.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformatio; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; Viral_RdRp; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CRYPTOCHROME_C; UNKNOWN_1.
DR Coar protein; Envelope protein; Glycoprotein; Nonstructural protein;
DR Polyprotein; Transmembrane.
DR SEQUENCE 3010 AA; 326857 MW; EA7D306A4BA2E224 CRC64;

Query Match 95.8%; Score 3034; DB 2; Length 3010;
Best Local Similarity 96.6%; Pred.No. 1.8e-227;
Matches 571; Conservative 14; Mismatches 6; Indels 0; Gaps 0;

Qy 5 SMSYTWTCALITPCAABESKLPINALNSLRHNLVSTTSRSASLRQKVTDFRLQVL 64
Db 2420 SMSYTWTCALITPCAABESKLPINPLNSLRHNLVSTTSRSASLRQKVTDFRLQVL 2479

Qy 65 DDHYRDVLKEMKAKASTVKALLSVEERACKLTPPHSAKSKFGYGAQVRSLSRAVNHIR 124
Db 2480 DDHYRDVLKEMKAKASTVKALLSVEERACKLTPPHSAKSKFGYGAQVRSLSRAVNHIR 2539

Qy 125 SVWKDLLEDTTPIOTTIMANNEVFCVQPEKGRKPARLIVFPDLGVRCCKMALYDVVS 184
Db 2540 SVWKDLLEDTTPIOTTIMANNEVFCVQPEKGRKPARLIVFPDLGVRCCKMALYDVVS 2599

Qy 185 TLPQAVMGSSYGFQYSPQORVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
Db 2600 TLPQAVMGSSYGFQYSPQORVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 2659

Qy 245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKQNGCRRASGVLTTCGNTLTCTYLKAA 304
Db 245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKQNGCRRASGVLTTCGNTLTCTYLKAA 304

```

```

Db 2660 QCCDLAPEARQAIRSLTERLYVGGPMTNSKQNGCRRASGVLTTCGNTLTCTYLKAA 2719
Qy 305 AACRAAKLQDCTMLVNGDDLVI CESAGTQBDAAASLRVFTAMTRYSAPPDPPOPEYDL 364
Db 2720 AACRAAKLQDCTMLVNGDDLVI CESAGTQBDAAASLRVFTAMTRYSAPPDPPOPEYDL 2779
Qy 365 ELITSCSNVSVAHDASGKRVYVYTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTLM 424
Db 2780 ELITSCSNVSVAHDASGKRVYVYTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTLM 2839
Qy 425 ARMLTMTHFFSILLAOEQLKALCOIYGACYSIEPLDLPQIERLHGLSFAFSLHSYSYG 484
Db 2840 ARMLTMTHFFSILLAOEQLKALCOIYGACYSIEPLDLPQIERLHGLSFAFSLHSYSYG 2899
Qy 485 EINRVASCLRLKLGVPPLRVWRHRSVRKLLSQGGRAAICGKYLFWNAVRTKLTPTIP 544
Db 2900 EINRVASCLRLKLGVPPLRVWRHRSVRKLLSQGGRAAICGKYLFWNAVRTKLTPTIP 2959
Qy 545 AASRLDLSGFWFVAGYGGDIYHSLSRAPRPFMLCLLLSVGVGYLLPNR 595
Db 2960 AASRLDLSGFWFVAGYGGDIYHSLSRAPRPFMLCLLLSVGVGYLLPNR 3010

RESULT 6
Q02829 PRELIMINARY; PRT; 3010 AA.
AC Q02829;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE Polypeptide.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HC-J4;
RX MEDLINE=92391112; PubMed=1325713;
RA Okamoto H., Kojima M., Okada S., Yoshizawa H., Iizuka H., Tanaka T.,
RT "Genetic drift of hepatitis C virus during an 8.2-year infection in a
RL chimpanzee: variability and stability."
RL Virology 190:894-899(1992).
DR EMBL; D13558; BAA02756.1; -.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSSP; Q8JYS1; 1CW.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformatio; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.

```



```
Db 2420 SMSYTWTCALITPCAABESKLPINPLNSLLRHHNMVYATTSRSAGLRQKKVTFDRLOVL 2479
QY 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
Db 2480 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVTHIR 2539
QY 125 SVWKDLLEDDTPIQTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 184
Db 2540 SVWKDLLEDDTETPISTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 2599
QY 185 TLPOAVMGSSYGFQYSPKORVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESTY 244
Db 2600 TLPOAVMGSSYGFQYSPKORVEFLVNTWKSKCPMGFSYDTRCFDSTVTENDIRVEESTY 2659
QY 245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNGYRCRCSAGVLTTCGNTLTCLYKAA 304
Db 2660 QCCDLAPEAKLAISLTERLYIGGPLTNSKGQNGYRCRCSAGVLTTCGNTLTCLYKAT 2719
QY 305 AACRAAKLODCTMLVNGDDLVIWICESAGTQEDAAASLRVFTTEAMTRYSPAPGDPPEYDL 364
Db 2720 AACRAAKLRDCTMLVNGDDLVIWICESAGTQEDAAASLRVFTTEAMTRYSPAPGDPPEYDL 2779
QY 365 ELITSCSSNVSAVHADSGKRVYVLTDPVPLARAAMETARHTPVNSWLGNIIMVAPTILW 424
Db 2780 ELITSCSSNVSAVHADSGKRVYVLTDPVPLARAAMETARHTPVNSWLGNIIMVAPTILW 2839
QY 425 ARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIRLHGLSAPLSHSYSPG 484
Db 2840 ARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIRLHGLSAPLSHSYSPG 2899
QY 485 EINRVASCLRLKGLVPPPLRVWRHARSVRAKLLSQGGRAAICGKYLFWNAVRTKLTPTIP 544
Db 2900 EINRVASCLRLKGLVPPPLRVWRHARSVRAKLLSQGGRAATCGKYLFWNAVRTKLTPTIP 2959
QY 545 AASRLDLSGFWAGYSGGDIYHLSRSPRPFMCLGILLSVGVGIYALPNR 595
Db 2960 AASRLDLSGFWAGYSGGDIYHLSRSPRPFMCLGILLSVGVGIYALPNR 3010

RESULT 8:
Q81825
ID Q81825 PRELIMINARY; PRT; 3010 AA.
AC Q81825;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE mRNA, complete cds.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OX Hepacivirus.
OX NCBI_TaxID=11103;
EN [1]
RP SEQUENCE FROM N.A.
RA Cho J.-M.;
RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
RL EMBL; M96362; AAA45721.1; -.
DR PIR; A61196; A61196.
DR PIR; PS0329; PS0329.
DR HSBF; Q81755; IDRP.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
```

```
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV NS5b.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept Ser Cys.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA pol BS PS.
DR InterPro; IPR007094; RNA pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; Viral RdRp; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN_1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
Polyprotein; Transmembrane.
SQ SEQUENCE 3010 AA; 326925 MW; PE997D54EE05142B CRC64;

Query Match 95.6%; Score 3027; DB 2; Length 3010;
Best Local Similarity 96.8%; Pred. No. 6.5e-227;
Matches 572; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 5 SMSYTWTCALITPCAABESKLPINALNSLLRHHNLVYSTTSRSASLRQKKVTFDRLOVL 64
Db 2420 SMSYTWTCALITPCAABESKLPINPLNSLLRHHNMVYATTSRSAGLRQKKVTFDRLOVL 2479
QY 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
Db 2480 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVTHIR 2539
QY 125 SVWKDLLEDDTPIQTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 184
Db 2540 SVWKDLLEDDTETPISTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 2599
QY 185 TLPOAVMGSSYGFQYSPKORVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESTY 244
Db 2600 TLPOAVMGSSYGFQYSPKORVEFLVNTWKSKCPMGFSYDTRCFDSTVTENDIRVEESTY 2659
QY 245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNGYRCRCSAGVLTTCGNTLTCLYKAA 304
Db 2660 QCCDLAPEAKLAISLTERLYIGGPLTNSKGQNGYRCRCSAGVLTTCGNTLTCLYKAT 2719
QY 305 AACRAAKLODCTMLVNGDDLVIWICESAGTQEDAAASLRVFTTEAMTRYSPAPGDPPEYDL 364
Db 2720 AACRAAKLRDCTMLVNGDDLVIWICESAGTQEDAAASLRVFTTEAMTRYSPAPGDPPEYDL 2779
QY 365 ELITSCSSNVSAVHADSGKRVYVLTDPVPLARAAMETARHTPVNSWLGNIIMVAPTILW 424
Db 2780 ELITSCSSNVSAVHADSGKRVYVLTDPVPLARAAMETARHTPVNSWLGNIIMVAPTILW 2839
QY 425 ARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIRLHGLSAPLSHSYSPG 484
Db 2840 ARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIRLHGLSAPLSHSYSPG 2899
QY 485 EINRVASCLRLKGLVPPPLRVWRHARSVRAKLLSQGGRAAICGKYLFWNAVRTKLTPTIP 544
Db 2900 EINRVASCLRLKGLVPPPLRVWRHARSVRAKLLSQGGRAATCGKYLFWNAVRTKLTPTIP 2959
```

QY 545 AASRLDLSGFWAGYSGGDIYHLSRPRWFMCLLLSVGVGIYLLPNR 595
 Db 2960 AASRLDLSGFWAGYSGGDIYHLSRPRWFMCLLLSVGVGIYLLPNR 3010
 RESULT 9
 Q9J3H0 PRELIMINARY; PRT; 3010 AA.
 ID Q9J3H0
 AC Q9J3H0;
 DT 01-OCT-2000 (TremBLrel. 15, Created)
 DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
 DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
 DE Polyprotein.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MD22;
 RA Nagayama K., Kurosaki M., Enomoto N., Miyaoka Y., Marumo F., Sato C.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF207763; AAF65953.1; -;
 DR PIR; A61196; A61196.
 DR PIR; P00246; P00246.
 DR PIR; P00254; P00254.
 DR PIR; P00329; P00329.
 DR HSSP; Q8UYS1; ICWX.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0005524; F:ATP-dependent helicase activity; IEA.
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR GO; GO:0019087; P:viral transformation; IEA.
 DR InterPro; IPR000345; CytC heme_BS.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH_N.
 DR InterPro; IPR002522; HCV_capsid.
 DR InterPro; IPR002521; HCV core.
 DR InterPro; IPR002519; HCV env.
 DR InterPro; IPR002531; HCV NS1.
 DR InterPro; IPR000745; HCV NS4a.
 DR InterPro; IPR001490; HCV NS4b.
 DR InterPro; IPR002868; HCV NS5a.
 DR InterPro; IPR002166; HCV NS5a.
 DR InterPro; IPR001650; Helicase_C.
 DR InterPro; IPR004109; Peptidase_S29.
 DR InterPro; IPR009003; Pept_Ser_Cys.
 DR InterPro; IPR002518; Pept_U39_HCV_NS2.
 DR InterPro; IPR007095; RNA_pol_DS_P5.
 DR InterPro; IPR007094; RNA_pol_Psivir.
 DR Pfam; PF01543; HCV_capsid; 1.
 DR Pfam; PF01542; HCV_core; 1.
 DR Pfam; PF01539; HCV_env; 1.
 DR Pfam; PF01560; HCV_NS1; 1.
 DR Pfam; PF01538; HCV_NS2; 1.
 DR Pfam; PF02907; HCV_NS3; 1.
 DR Pfam; PF01006; HCV_NS4a; 1.
 DR Pfam; PF01001; HCV_NS4b; 1.
 DR Pfam; PF01506; HCV_NS5a; 1.
 DR Pfam; PF00271; Helicase_C; 1.
 DR Pfam; PF00398; Viral_RdRp; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 Polyprotein; Transmembrane.

SQ SEQUENCE 3010 AA; 327098 MW; 737EEF31E3C2B28D CRC64;
 Query Match 95.5%; Score 3026; DB 2; Length 3010;
 Best Local Similarity 96.3%; Pred. No. 7.8e-227;
 Matches 569; Conservative 14; Mismatches 8; Indels 0; Gaps 0;
 QY 5 SMSYTWGALITPCAAEESKLPINALSNLRLHNLVYSTTSRASLRQKKVTFDRQLQVL 64
 Db 2420 SMSYTWGALITPCAAEESKLPINALSNLRLHNLVYSTTSRASLRQKKVTFDRQLQVL 2479
 QY 65 DDHYRDVLEKMKAKASTVAKLLSVEACKLTPPHSAKSGFYGAKDVRSLSSRAVNHIR 124
 Db 2480 DDHYRDVLEKMKAKASTVAKLLSVEACKLTPPHSAKSGFYGAKDVRSLSSRAVNHIR 2539
 QY 125 SVWDLLEDTDTPITQTIMAKNEVFCVQPEKGGKPARLIVFPDLGVRCVKMALYDVVS 184
 Db 2540 SVWDLLEDTDTPITQTIMAKNEVFCVQPEKGGKPARLIVFPDLGVRCVKMALYDVVS 2599
 QY 185 TLPOAVMGSSVGYFOYSPKORVEPLVNTWKAACKPMGFSDYTRCFDSTVTENDIRVESIY 244
 Db 2600 TLPOAVMGSSVGYFOYSPKORVEPLVNTWKAACKPMGFSDYTRCFDSTVTENDIRVESIY 2659
 QY 245 QCCLAPPEARQAINSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTSCGNLTLCYLKAA 304
 Db 2660 QCCLAPPEARQAINSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTSCGNLTLCYLKAS 2719
 QY 305 AACRAAKLQDCTMLVNGDDLWICESAGTOEDASIRVFTTEMTRYSAAPPDPPQPEYDL 364
 Db 2720 AACRAAKLQDCTMLVNGDDLWICESAGTOEDASIRVFTTEMTRYSAAPPDPPQPEYDL 2779
 QY 365 ELITSCSNVSVAHDSAGKRVVYLTDRPTVPLARAAMETARHTPVNSWLGNIIMYAPTLW 424
 Db 2780 ELITSCSNVSVAHDSAGKRVVYLTDRPTVPLARAAMETARHTPVNSWLGNIIMYAPTLW 2839
 QY 425 ARMLTMHTFSSILIAQOEKALDCQIYGACYSIEPLDLPOIIRLHGLSAFSLHSYSPG 484
 Db 2840 ARMLTMHTFSSILIAQOEKALDCQIYGACYSIEPLDLPOIIRLHGLSAFSLHSYSPG 2899
 QY 485 EINRVASCLRLKGLVPLRVWRHRSVRKLLSOGGAAATCGKYLEFNWAVTKLKLTPIP 544
 Db 2900 EINRVASCLRLKGLVPLRVWRHRSVRKLLSOGGAAATCGKYLEFNWAVTKLKLTPIP 2959
 QY 545 AASRLDLSGFWAGYSGGDIYHLSRPRWFMCLLLSVGVGIYLLPNR 595
 Db 2960 AASRLDLSGFWAGYSGGDIYHLSRPRWFMCLLLSVGVGIYLLPNR 3010
 RESULT 10
 Q9QIX6 PRELIMINARY; PRT; 3010 AA.
 ID Q9QIX6
 AC Q9QIX6;
 DT 01-MAY-2000 (TremBLrel. 13, Created)
 DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
 DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
 DE Polyprotein.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MD8-1;
 RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyaoka Y.,
 RA Tazawa J.I., Izumi N., Marumo F., Sato C.;
 RT "Time-related changes in full-length hepatitis C virus and hepatitis
 activity";
 RT Virology 263:244-253 (1999).
 DR EMBL; AF165059; AAD56194.1; -;
 DR PIR; A61196; A61196.
 DR PIR; P00246; P00246.
 DR PIR; P00254; P00254.
 DR PIR; P00804; P00804.

```
DR PIR; PS0329; PS0329.
DR HSSP; Q8JYS1; 1CW.
DR MEROPS; S29.002; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR00345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4b.
DR InterPro; IPR001490; HCV NS5a.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV NS5a.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA pol_DS_PS.
DR InterPro; IPR007094; RNA pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; Viral RdRp; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
Polyprotein; Transmembrane.
SQ SEQUENCE 3010 AA; 327468 MW; 4613744EC4D4A013 CRC64;

Query Match          95.5%; Score 3025; DB 2; Length 3010;
Best Local Similarity 96.4%; Pred. No. 9.3e-227;
Matches 570; Conservative 15; Mismatches 6; Indels 0; Gaps 0;

QY 5 SMSYTWTCALITPCAABESKLPINALNSLRHNLVYSTSRASLRQKKVTFDRQLVL 64
Db 2420 SMSYTWTCALITPCAABESKLPINPLNSLRHNLVYATTSRASLRQKKVTFDRQLVL 2479

QY 65 DDHYRDVLKEMKAKASTYKALLSVEACKLTPPHSAKSKFGYKADYRSLSSRAVNHIR 124
Db 2480 DDHYRDVLKEMKAKASTYKALLSVEACKLTPPHSAKSKFGYKADYRSLSSRAVNHIR 2539

QY 125 SVWKDLLEDTDPIOTTIMAKNEVFCVPEKGRKPABLI VFPDILGVRVCEKALYDVVS 184
Db 2540 SVWKDLLEDTEPIOTTIMAKSEVFCVPEKGRKPABLI VFPDILGVRVCEKALYDVVS 2599

QY 185 TLPAQVMSGYGFQYSPKQRFVFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIIY 244
Db 2600 TLPAQVMSGYGFQYSPKQRFVFLVNTWKSKCPMGFSYDTRCFDSTVTESDIRTEESIIY 2659

QY 245 QCCDLAPARQAIRSLTERLYVGGPMWNSKGNCCGYRRCRASGLVTTSCGNLTLCYLKAA 304
Db 2660 QCCDLAPARQAIRSLTERLYVGGPMWNSKGNCCGYRRCRASGLVTTSCGNLTLCYLKAS 2719
```

```
QY 305 AACRAAKLQDCTMLVNGDDLVLVVCESAGTQEDASLRVFTTEAMTRYSPGPPQPEYDL 364
Db 2720 AACRAAKLQDCTMLVNGDDLVLVVCESAGTQEDASLRVFTTEAMTRYSPGPPQPEYDL 2779

QY 365 ELITSCSSNVSAHDASGRVYLLTRDTPVPLARAAMETARHTPNSWLGNIIMYAPTLM 424
Db 2780 ELITSCSSNVSAHDASGRVYLLTRDTPVPLARAAMETARHTPNSWLGNIIMYAPTLM 2839

QY 425 ARMILMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPOIIEHLGLSAFSLHSYSPG 484
Db 2840 ARMILMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPOIIEHLGLSAFSLHSYSPG 2899

QY 485 EINRVASCLRLKLGVPPLRVWRHRRARSVRKALLSQSGRAAIICGYLFNNAVRTKLTPTIP 544
Db 2900 EINRVASCLRLKLGVPPLRVWRHRRARSVRKALLSQSGRAATCGKYLEFNNAVRTKLTPTIP 2959

QY 545 AASRLDLSCGWFVAGYSGGDIYHSLSRARPRWFMCLLLLSVGVGYLLPNR 595
Db 2960 AASRLDLSCGWFVAGYSGGDIYHSLSRARPRWFMCLLLLSVGVGYLLPNR 3010

RESULT 11
QSQIX5 PRELIMINARY; PRT; 3010 AA.
AC Q9QIX5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=MD8-2;
RX MEDLINE=20013325; PubMed=10544098; DOI=10.1006/viro.1999.9924;
RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyasaka Y.,
RA Tazawa J., Izumi N., Marumo F., Sato C.;
RT "Time-related changes in full-length hepatitis C virus and hepatitis
RT activity.";
RL Virology 263:244-253(1999).
DR EMRL; AF165060; AAD56195.1; -.
DR PIR; A61196; A61196.
DR PIR; P00246; P00246.
DR PIR; P00254; P00254.
DR PIR; P00804; P00804.
DR PIR; PS0329; PS0329.
DR HSSP; Q8JYS1; 1CW.
DR MEROPS; S29.002; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR00345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4b.
DR InterPro; IPR001490; HCV NS5a.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV NS5a.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA pol_DS_PS.
DR InterPro; IPR007094; RNA pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; Viral RdRp; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
```

```

DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; Viral_RdRP; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3010 AA; 327295 MW; 8B99F1EBAG6A50F56 CRC64;

Query Match 95.5%; Score 3024; DB 2; Length 3010;
Blast Local Similarity 96.4%; Pred.No.1.le-226; Mismatches 7; Indels 0; Gaps 0;
Matches 570; Conservative 14;

QY 5 SMSYTWTCALITPCAABESKLPINALNSLRHNLVYSTTSRGAASLRQKQKVTDRLOVL 64
DB 2420 SMSYTWTCALITPCAABESKLPINPLNSLRHNLVYSTTSRGAASLRQKQKVTDRLOVL 2479
QY 65 DDHYRDVLKEMKAKASTVAKLLSVEACKLTPPHSAKSKFGYGAKVRSLSRAVNHR 124
DB 2480 DDHYRDVLKEMKAKASTVAKLLSVEACKLTPPHSAKSKFGYGAKVRSLSRAVNHR 2539
QY 125 SVWDLLEDTTPIOTTIMANNEVCPOPEKGGKPARLIYFPDLGVRCFKMALYDVVS 184
DB 2540 SVWDLLEDTTPIOTTIMANNEVCPOPEKGGKPARLIYFPDLGVRCFKMALYDVVS 2599
QY 185 TLPQAVMGSSYGFQYSPQRVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
DB 2600 TLPQAVMGSSYGFQYSPQRVEFLVNTWKAKCPMGFSYDTRCFDSTVTESDIRTEESIY 2659
QY 245 QCCLDAPARQAIRSLTERLYVGGPMTNKSGNCGYRRCRASGVLTTSCTGNTLTCYLKAA 304
DB 2660 QCCLDAPARQAIRSLTERLYVGGPMTNKSGNCGYRRCRASGVLTTSCTGNTLTCYLKAT 2719
QY 305 AACRAAKLQDCTMLVNGDGLVVICESACTOEDASLRVFTAMTRYSAAPPDPPQPEYDL 364
DB 2720 AACRAAKLQDCTMLVNGDGLVVICESACTOEDASLRVFTAMTRYSAAPPDPPQPEYDL 2779
QY 365 ELITSCSSNVSAHDASGRVYVLTTRDPTVPLARAAMETARHTPVNSWLGNIIMVAPTLW 424
DB 2780 ELITSCSSNVSAHDASGRVYVLTTRDPTVPLARAAMETARHTPVNSWLGNIIMVAPTLW 2839
QY 425 ARMLMTHTFFSILLAOEQLKALDCQIYGACYSIEPLDLPQIERLHGLSAFSLHSYSPG 484
DB 2840 ARMLMTHTFFSILLAOEQLKALDCQIYGACYSIEPLDLPQIERLHGLSAFSLHSYSPG 2899
QY 485 EINRVASCLRLKGLVPLRWHRARSRVAKLLSQGGRAAI CGKLYFNWAVTKLKLTP 544
DB 2900 EINRVASCLRLKGLVPLRWHRARSRVAKLLSQGGRAATCGKLYFNWAVTKLKLTP 2959
QY 545 AASRLDLSGWFGVAGSGGDIYHLSRARPRFMWCLLLLSVGVGYLLLPNR 595
DB 2960 AASRLDLSGWFGVAGSGGDIYHLSRARPRFMWCLLLLSVGVGYLLLPNR 3010

RESULT 12
ID P89966 PRELIMINARY; PRT; 3010 AA.

```

```

AC P89966;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DE RNA for polyprotein, complete cds.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=type 1b;
RA TANAKA T.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=type 1b;
RA Tanaka T.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; D89872; BAA14035.1; -.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSP; O81755; 1DXP.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003958; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; Viral_RdRP; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3010 AA; 327034 MW; E075BD9C9FD8D1261 CRC64;

Query Match 95.4%; Score 3022; DB 2; Length 3010;

```



```
Best Local Similarity 96.4%; Pred. No. 1.6e-226;
Matches 570; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY 5 SMSYTWGALITPCAAEESKLPINALSNLLRHHNLVSTTSRSASLRQKKVTFDRLOVL 64
D 2420 SMSYTWGALITPCAAEESKLPINPLSNLLRHHNSVSTTSRSASLRQKKVTFDRLOVL 2479
QY 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNH 124
D 2480 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNH 2539
QY 125 SVWKDLLEDTPDTPIOTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKALYDVVS 184
D 2540 SVWKDLLEDTPDTPIOTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKALYDVVS 2599
QY 185 TLPOAVMGSSYGFQSPQORVFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
D 2600 TLPOAVMGSSYGFQSPQORVFLVNTWKSKCPMGFSYDTRCFDSTVTENDIRVEESIY 2659
QY 245 QCCDLAPBARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTCGNTLTCLYKAA 304
D 2660 QCCDLAPBARQAIRSLTERLYIGGPLTNSKQNGCYRRCRASGVLTTCGNTLTCLYKAT 2719
QY 305 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPPEYDL 364
D 2720 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPPEYDL 2779
QY 365 ELITSCSSNVSAHDAHGKRVYLLTRDPTVPLARAAMETARHTPVNSWLGNIIMTAPLW 424
D 2780 ELITSCSSNVSAHDAHGKRVYLLTRDPTVPLARAAMETARHTPVNSWLGNIIMTAPLW 2839
QY 425 ARMILMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
D 2840 ARMILMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2899
QY 485 EIRNVASCLRKLGVPPLRWRHRAARSRAKLLSQGGRAAI CKYLPNVAWRTKLTIP 544
D 2900 EIRNVASCLRKLGVPPLRWRHRAARSRAKLLSQGGRAATCKYLPNVAWRTKLTIP 2959
QY 545 AASRLDLSGFWVAGYSGGDIYHLSRARPRWFMCLLLLSVGVGYLLPNR 595
D 2960 AASRLDLSGFWVAGYSGGDIYHLSRARPRWFMCLLLLSVGVGYLLPNR 3010

RESULT 13
QSDTE0 PRELIMINARY; PRT; 3014 AA.
AC QSDTE0;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI TaxID=111103;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Serum;
RA Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,
RA Hatahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijikata M.,
RA Mishiro S.;
RT "Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients
RT with hepatocellular carcinoma; the 'progression score' revisited.";
RL Hepatol. Res. 20:161-171(2001).
DR EMBL; AB049097; BAB18810.1; -.
DR PIR; A61196; A61196.
DR PIR; PS0329; PS0329.
DR HSSP; Q81755; 1DXP.
DR MEROPS; S29.002; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
```

```
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; Viral RdRP; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3014 AA; 327498 MW; ED9BF4F94BDB6287 CRC64;

Query Match 95.4%; Score 3022; DB 2; Length 3014;
Best Local Similarity 96.1%; Pred. No. 1.6e-226;
Matches 568; Conservative 15; Mismatches 8; Indels 0; Gaps 0;

QY 5 SMSYTWGALITPCAAEESKLPINALSNLLRHHNLVSTTSRSASLRQKKVTFDRLOVL 64
D 2424 SMSYTWGALITPCAAEESKLPINPLSNLLRHHNMIYATTSRSAGLRQKKVTFDRLOVL 2483
QY 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNH 124
D 2484 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNH 2543
QY 125 SVWKDLLEDTPDTPIOTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKALYDVVS 184
D 2544 SVWKDLLEDTPDTPIOTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKALYDVVS 2603
QY 185 TLPOAVMGSSYGFQSPQORVFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
D 2604 TLPOAVMGSSYGFQSPQORVFLVNTWKSKCPMGFSYDTRCFDSTVTENDIRVEESIY 2663
QY 245 QCCDLAPBARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTCGNTLTCLYKAA 304
D 2664 QCCDLAPBARQAIRSLTERLYIGGPLTNSKQNGCYRRCRASGVLTTCGNTLTCLYKAS 2723
QY 305 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPPEYDL 364
D 2724 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPPEYDL 2783
QY 365 ELITSCSSNVSAHDAHGKRVYLLTRDPTVPLARAAMETARHTPVNSWLGNIIMTAPLW 424
```

Db 2784 ELITSCSSNVSAHDASGRVYLTDRPTTPARAAMETARHTPVNSWLGNIIMVAPTLW 2843
Qy 425 ARMLMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLPQIIERLHGLSFAFSLHSYSPG 484
Db 2844 ARMLMTHFFSILLVQELKALDCQIYGACYSIEPLDLPQIIERLHGLSFAFSLHSYSPG 2903
Qy 485 EINRVASCLRLKGLVPPPLRVWRHRRARSVRAKLLSQGGRAAICGKYLEFNWAVRTKLTPIP 544
Db 2904 EINRVASCLRLKGLVPPPLRVWRHRRARSVRAKLLSQGGRAATCGKYLEFNWAVRTKLTPIP 2963
Qy 545 AASRLDLSGMFVAGYSGGDIYHLSLRARPRFWMCLLLLSVGVGIYLLPNR 595
Db 2964 AASRLDLSGMFVAGYSGGDIYHLSVRARPRFWMCLLLLSVGVGIYLLPNR 3014

RESULT 14
CO09796
ID CO09796 PRELIMINARY; PRT; 3010 AA.
AC CO09796;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Genomic RNA, complete cds.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=genotype II/1b;
RX MEDLINE=97170750; PubMed=9018054;
RA Sugiyama K., Kato N., Mizutani T., Ikeda M., Tanaka T., Shimotohno K.;
RT "Genetic analysis of the hepatitis C virus (HCV) genome from HCV-
RT infected human T cells."
RL J. Gen. Virol. 78:329-336(1997).
DR EMBL; D85516; BAA19625.1; -.
DR PIR; A61196; A61196.
DR PIR; F00804; F00804.
DR PIR; F50329; F50329.
DR HSSP; Q8JYS1; ICWX.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0006508; F: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0019087; P: viral genome replication; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RDRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_PS.
DR InterPro; IPR007094; RNA_pol_Psivir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV_core; 1.

Pfam; PF01539; HCV env; 1.
Pfam; PF01560; HCV_NS1; 1.
Pfam; PF01538; HCV_NS2; 1.
Pfam; PF02907; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; Helicase_C; 1.
Pfam; PF00998; Viral_RDRP; 1.
SMART; SM00487; DEAD_C; 1.
PROSITE; PS00190; CYTOCHROME_C; UNKNOWN 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3010 AA; 327266 MW; 4324643E6A1D5CE4 CRC64;

Query Match 95.4%; Score 3021; DB 2; Length 3010;
Best Local Similarity 96.3%; Pred. No. 1.9e-226;
Matches 569; Conservative 14; Mismatches 8; Indels 0; Gaps 0;

Qy 5 SMSVTWTGALITPCAABESKLPINALSNSLLRHNLVYVTSRSASLRQKKVTFDRLOVL 64
Db 2420 SMSVTWTGALITPCAABESKLPINPLSNSLLRHNLVYVTSRSAGLRQKKVTFDRLOVL 2479
Qy 65 DDHYRDVLKEMKAKASTVKAKLLSVEBACKLTPPHSAKSKFGYGAQDVRSLSRAVNHIR 124
Db 2480 DDHYRDVLKEMKAKASTVKAKLLSIEBACKLTPPHSAKSKYGYGAQDVRLSSRAVNHIR 2539
Qy 125 SVWKDLLEDTDTPQTITIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 184
Db 2540 SVWKDLLEDTDTPITIMAKSEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 2599
Qy 185 TLPQAVNGSSVGFQYSPKORVEPLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVESIY 244
Db 2600 TLPQAVNGSSVGFQYSPAQRVEFLVNTWKSCKCPMGFSYDTRCFDSTVTENDIRVESIY 2659
Qy 245 QCCLDAPEARQAIRSLTERLYVGGPMNTNSKQNGCYRRCRASGVLTTSCGNLTLCYLKAA 304
Db 2660 QCCLDAPEAKAIRSLTERLYVGGPLTNSKQNGCYRRCRASGVLTTSCGNLTLCYLKAS 2719
Qy 305 AACRAAKLQDCTMLVNGDDLVCESAGTQEDDAASLRVFTFAMTRYSAPODPPQPEYDL 364
Db 2720 AACRAAKLQDCTMLVNGDDLVCESAGTQEDDAASLRVFTFAMTRYSAPODPPQPEYDL 2779
Qy 365 ELITSCSSNVSAHDASGRVYLTDRPTTPARAAMETARHTPVNSWLGNIIMVAPTLW 424
Db 2780 ELITSCSSNVSAHDASGRVYLTDRPTTPARAAMETARHTPVNSWLGNIIMVAPTLW 2839
Qy 425 ARMLMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLPQIIERLHGLSFAFSLHSYSPG 484
Db 2840 ARMLMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLPQIIERLHGLSFAFSLHSYSPG 2899
Qy 485 EINRVASCLRLKGLVPPPLRVWRHRRARSVRAKLLSQGGRAAICGKYLEFNWAVRTKLTPIP 544
Db 2900 EINRVASCLRLKGLVPPPLRVWRHRRARSVRAKLLSQGGRAATCGKYLEFNWAVRTKLTPIP 2959
Qy 545 AASRLDLSGMFVAGYSGGDIYHLSLRARPRFWMCLLLLSVGVGIYLLPNR 595
Db 2960 AASRLDLSGMFVAGYSGGDIYHLSLRARPRFWMCLLLLSVGVGIYLLPNR 3010

RESULT 15
CO09930
ID CO09930 PRELIMINARY; PRT; 591 AA.
AC CO09930;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RNA-dependent RNA-polymerase (Fragment).
GN Name=NS5b;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;


```

[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=98001363; PubMed=9343198;
RA Lohmann V., Korner F., Herian U., Bartenschlager R.;
RT "Biochemical properties of hepatitis C virus NS5B RNA-dependent RNA
RT polymerase and identification of amino acid sequence motifs essential
RT for enzymatic activity.";
RL J. Virol. 71:8416-8428(1997).
DR EMBL; Z97730; CAB10747.1; -.
DR PDB; INHU; X-ray; A/B=1-570.
DR PDB; INHV; X-ray; A/B=1-570.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF00998; Viral_RdRp; 1.
FT NON_TER 1
SQ SEQUENCE 591 AA; 65753 MW; 6FA9C2F970260349 CRC64;

Query Match 95.4%; Score 3020; DB 2; Length 591;
Best Local Similarity 96.6%; Pred.No. 2.5e-227;
Matches 571; Conservative 12; Mismatches 8; Indels 0; Gaps 0;

Qy 5 SMSYTWGALITPCAERSKLPINALSNLSLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 1 SMSYTWGALITPCAERSKLPINALSNLSLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 60
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 65 DDHYRDVLKEMKAKASTVKALLSVEEACKLTPPHSAKSKFGYGAQVRSLSRAVNHIR 124
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 61 DDHYRDVLKEMKAKASTVKALLSVEEACKLTPPHSAKSKFGYGAQVRSLSRAVNHIR 120
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 125 SWKDLLEDTPPTQTTIMAKNEVFCQPEKGRKPARLIVPDDLGVRVCEKMALYDVWS 184
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 121 SWKDLLEDTPPTQTTIMAKNEVFCQPEKGRKPARLIVPDDLGVRVCEKMALYDVWS 180
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 185 TLPOAVMGSSYGVFOYSPKQRYVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVESIY 244
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 181 TLPOAVMGSSYGVFOYSPKQRYVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVESIY 240
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKQNCQYRCRASGLVITSCGNLTLCYLKAA 304
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 241 QCCDLAPEARQAIRSLTERLYVGGPMTNSKQNCQYRCRASGLVITSCGNLTLCYLKAS 300
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 305 AACRAAKLQDCTMLVNGDDLVI CESAGTQBDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 301 AACRAAKLQDCTMLVCGDDLVI CESAGTQBDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 360
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 365 ELITSCSSNVSVADHASKRYVYLTROPTVPLARAAMETABHTPVNSWLGNIIMYAPTLW 424
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 361 ELITSCSSNVSVADHASKRYVYLTROPTVPLARAAMETABHTPVNSWLGNIIMYAPTLW 420
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 425 ARMILMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIERLHGLSAFSLHSYSPG 484
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 421 ARMILMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIERLHGLSAFSLHSYSPG 480
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 485 EINRVASCLRLKGVPPPLRVWRHRSVRRAKLLSQGGRAAICGKYLEFNWVTKLKTPIP 544
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 481 EINRVASCLRLKGVPPPLRVWRHRSVRRAKLLSQGGRAAICGKYLEFNWVTKLKTPIP 540
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 545 AASRLDLSGWFPVAGYSGGDIYVHSLSRAPRPFMLCLLLLSVGVIYLLPNR 595
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 541 AASQLDLSGWFPVAGYSGGDIYVHSLSRAPRPFMFMCCLLLLSVGVIYLLPNR 591
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2005, 14:39:31 ; Search time 27 Seconds
(without alignments)
1661.633 Million cell updates/sec

Title: US-10-712-479-2

Perfect score: 3167

Sequence: 1 MASMSMTWTGALITPCA.....LLSVGVGIYLLPNRHHHHH 601

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3016	95.2	591	3	US-08-952-981A-1
2	3016	95.2	997	1	US-08-324-977-50
3	3016	95.2	997	2	US-08-384-616-50
4	3016	95.2	997	2	US-08-904-686A-50
5	3016	95.2	997	3	US-09-315-850-50
6	3016	95.2	2201	3	US-08-952-981A-2
7	3016	95.2	2621	1	US-08-324-977-36
8	3016	95.2	2621	2	US-08-384-616-36
9	3016	95.2	2621	3	US-08-904-686A-36
10	3016	95.2	2621	3	US-09-315-850-36
11	3016	95.2	3010	1	US-08-324-977-2
12	3016	95.2	3010	1	US-08-324-977-14
13	3016	95.2	3010	2	US-08-384-616-2
14	3016	95.2	3010	2	US-08-384-616-14
15	3016	95.2	3010	2	US-08-904-686A-2
16	3016	95.2	3010	2	US-08-904-686A-14
17	3016	95.2	3010	3	US-09-315-850-2
18	3016	95.2	3010	3	US-09-315-850-14
19	3016	95.2	3010	3	US-09-014-416-3
20	3014	95.2	593	4	US-09-597-877-23
21	3011	95.1	621	4	US-09-898-297-1
22	3011	95.1	621	4	US-09-995-099-1
23	3011	95.1	621	4	US-10-238-282-1
24	3011	95.1	2620	1	US-08-324-977-32
25	3011	95.1	2620	2	US-08-384-616-32
26	3011	95.1	2620	2	US-08-904-686A-32
27	3011	95.1	2620	3	US-09-315-850-32

28	3005	94.9	1985	4	US-09-539-601-9	Sequence 9, Appli
29	3005	94.9	1985	4	US-09-539-601-12	Sequence 12, Appl
30	3005	94.9	1985	4	US-09-539-601-24	Sequence 24, Appl
31	3005	94.9	2201	4	US-09-539-601-6	Sequence 6, Appli
32	3005	94.9	2201	4	US-09-539-601-15	Sequence 15, Appl
33	3005	94.9	2201	4	US-10-029-907-3	Sequence 3, Appli
34	3005	94.9	3010	4	US-09-539-601-3	Sequence 3, Appli
35	3005	94.9	3010	4	US-09-539-601-27	Sequence 27, Appl
36	3004	94.9	1985	4	US-09-539-601-18	Sequence 18, Appl
37	3004	94.9	3010	4	US-09-539-601-21	Sequence 21, Appl
38	3002	94.8	1985	4	US-09-539-601-30	Sequence 30, Appl
39	3002	94.8	3010	4	US-09-539-601-33	Sequence 33, Appl
40	2993	94.5	591	4	US-09-720-095A-2	Sequence 2, Appli
41	2987	94.3	593	4	US-09-597-877-12	Sequence 12, Appl
42	2927	92.4	590	4	US-09-194-949A-15	Sequence 15, Appl
43	2879	90.9	576	4	US-09-541-990A-1	Sequence 1, Appli
44	2786	88.0	591	3	US-09-208-140-2	Sequence 2, Appli
45	2774	87.6	3011	3	US-09-014-416-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-08-952-981A-1
; Sequence 1, Application US/08952981A
; Patent No. 6383768
; GENERAL INFORMATION:
; APPLICANT: DE FRANCESCO, Raffaele
; APPLICANT: TOMELI, Licia
; APPLICANT: BEHRENS, Sven-Erik
; TITLE OF INVENTION: METHOD FOR REPRODUCING IN VITRO THE RNA-DEPENDENT RNA
; TITLE OF INVENTION: POLYMERASE AND TERMINAL NUCLEOTIDYL TRANSFERASE
; TITLE OF INVENTION: ACTIVITIES ENCODED BY HEPATITIS C VIRUS (HCV)
; FILE REFERENCE: IT0002P
; CURRENT APPLICATION NUMBER: US/08/952,981A
; CURRENT FILING DATE: 1998-03-23
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-08-952-981A-1

Query Match	95.2%	Score	3016	DB	3	Length	591
Best Local Similarity	96.6%	Pred. No.	2e-298				
Matches	571	Conservative	11	Mismatches	9	Indels	0
Gaps	0						
Qy	5	SMSYTWGALITPCAABESKLPINALNSLRHNLVYVSTTSRGSASLRQKVTFRDLQVL	64				
Db	1	SMSYTWGALITPCAABESKLPINALNSLRHNLVYVSTTSRGSASLRQKVTFRDLQVL	60				
Qy	65	DDHYRDLVKEMKAKASTVAKLLSVEEACKLTPPHSAKSKFGYGAOVRSLSRAVNHHR	124				
Db	61	DDHYRDLVKEMKAKASTVAKLLSVEEACKLTPPHSAKSKFGYGAOVRSLSRAVNHHR	120				
Qy	125	SVWKDLLEDTTPTQTTIMAKNEVFCVQPEKGGKPKARLIIVFPDLGVRVCEKMAIYDVVS	184				
Db	121	SVWKDLLEDTTPTQTTIMAKNEVFCVQPEKGGKPKARLIIVFPDLGVRVCEKMAIYDVVS	180				
Qy	185	TLPOAVNGSSYGFQYSPKORVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIIY	244				
Db	181	TLPOAVNGSSYGFQYSPKORVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIIY	240				
Qy	245	QCCDLAPPEARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGLVLTSCGNTLTCLYKAA	304				
Db	241	QCCDLAPPEARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGLVLTSCGNTLTCLYKAS	300				
Qy	305	AACRAAKLQDCTMLVNGDDLVIICESAGTQDAASLRVFTTEAMTRYSPAPGPPQPEYDL	364				
Db	301	AACRAAKLQDCTMLVNGDDLVIICESAGTQDAASLRVFTTEAMTRYSPAPGPPQPEYDL	360				

365	Qy	ELITCSSNVSAHDA	SGRVYVYTRDPT	PLAARAAWETARHTP	SVNLGNIIMYAPTLL	424
361	Db	ELITCSSNVSAHDA	SGRVYVYTRDPT	PLAARAAWETARHTP	SVNLGNIIMYAPTLL	420
425	Qy	ARMIWMTHFFS	LILAQOELEKAL	COIYGACVSTIEP	LDLPQIETRLHGLSAFSLHSYS	484
421	Db	ARMIWMTHFFS	LILAQOELEKAL	COIYGACVSTIEP	LDLPQIETRLHGLSAFSLHSYS	480
485	Qy	EINRVASCLRL	KLGVPPPLVRW	HRARSVRAKLLS	OGGRAAICGKYLFWNAVTKLKTIP	544
481	Db	EINRVASCLRL	KLGVPPPLVRW	HRARSVPARLL	SGGRAAICGKYLFWNAVTKLKTIP	540
545	Qy	AASRLDLSGFW	VAGVSGGDIYHS	LSRAPRPFMIL	CLLLLSVGVIYLLPNR	595
541	Db	AASRLDLSGFW	VAGVSGGDIYHS	LSRAPRPFMIL	CLLLLSVGVIYLLPNR	591

RESULT 2

```

US-08-324-977-50
; Sequence 50, Application US/08324977
; Patent No. 5747339
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westernman, Hattori, McLeeland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006

```

```
;
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
;
```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/324,977
FILING DATE: 18-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Stevens-Smith, Theresa M.
REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: 900703D
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
TELEX: 440142
INFORMATION FOR SEC ID NO: 50:

```



```

; MOLECULE TYPE: protein
US-08-904-686A-50

Query Match      95.2%; Score 3016; DB 2; Length 997;
Best Local Similarity 96.6%; Pred. No. 4.8e-298;
Matches 571; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

Qy 5 SMSYTWGALITPCAAEESKLPINALNSLRHNLVYSTTSRSASLRQKVTDFRLQVL 64
Db 407 SMSYTWGALITPCAAEESKLPINALNSLRHNLVYSTTSRSASLRQKVTDFRLQVL 466

Qy 65 DDHYRDVLKEMKAKASTVKALLSVEEACKLTPPHSAKSKFGYKADKVRSLSSRAVNHIR 124
Db 467 DDHYRDVLKEMKAKASTVKALLSVEEACKLTPPHSAKSKFGYKADKVRSLSSRAVNHIIH 526

Qy 125 SVWKDLLEDVTPIDTTIMAKNEVFCVQPEKGRKPARLI VFPDLGVRVCEKMALYDVVS 184
Db 527 SVWKDLLEDVTPIDTTIMAKNEVFCVQPEKGRKPARLI VFPDLGVRVCEKMALYDVVS 586

Qy 185 TLPQAVMGSSYGFQYSPKQRFVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVESIY 244
Db 587 TLPQAVMGSSYGFQYSPKQRFVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVESIY 646

Qy 245 QCCLDLAPARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGLVTTSCGNLTLCYLKAA 304
Db 647 QCCLDLAPARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGLVTTSCGNLTLCYLKAS 706

Qy 305 AACRAAKLQDCTMLVNGDDLWVICESAGTQBDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
Db 707 AACRAAKLQDCTMLVNGDDLWVICESAGTQBDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 766

Qy 365 ELITSCSSNSVSAHDASGKRVYLTDRPTVPLARAWEETARHTPVNSWMLGNIIMYAPTLW 424
Db 767 ELITSCSSNSVSAHDASGKRVYLTDRPTVPLARAWEETARHTPVNSWMLGNIIMYAPTLW 826

Qy 425 ARMLTHTFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIERLHGLSAFSLHSYSPG 484
Db 827 ARMLTHTFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIERLHGLSAFSLHSYSPG 886

Qy 485 EBNRVASCLRLGVPPLRWHRARSAVRALLSQGGRAAICGKYLFWNAVTKLKLTPIP 544
Db 887 EBNRVASCLRLGVPPLRWHRARSAVRALLSQGGRAAICGKYLFWNAVTKLKLTPIP 946

Qy 545 AASRLDLGWFVAGSGGDIYHLSLRAPRPFMLCLLLSVGVGYLLPNR 595
Db 947 AASRLDLGWFVAGSGGDIYHLSLRAPRPFMLCLLLSVGVGYLLPNR 997

RESULT 5
US-09-315-850-50
; Sequence 50, Application US/09315850
; Patent No. 6217872
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Ieao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeeland &
; ADDRESS: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0

```

```

; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/315,850
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686
; FILING DATE: 01-AUG-1997
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: McLeeland, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 900703G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 997 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-315-850-50

Query Match      95.2%; Score 3016; DB 3; Length 997;
Best Local Similarity 96.6%; Pred. No. 4.8e-298;
Matches 571; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

Qy 5 SMSYTWGALITPCAAEESKLPINALNSLRHNLVYSTTSRSASLRQKVTDFRLQVL 64
Db 407 SMSYTWGALITPCAAEESKLPINALNSLRHNLVYSTTSRSASLRQKVTDFRLQVL 466

Qy 65 DDHYRDVLKEMKAKASTVKALLSVEEACKLTPPHSAKSKFGYKADKVRSLSSRAVNHIR 124
Db 467 DDHYRDVLKEMKAKASTVKALLSVEEACKLTPPHSAKSKFGYKADKVRSLSSRAVNHIIH 526

Qy 125 SVWKDLLEDVTPIDTTIMAKNEVFCVQPEKGRKPARLI VFPDLGVRVCEKMALYDVVS 184
Db 527 SVWKDLLEDVTPIDTTIMAKNEVFCVQPEKGRKPARLI VFPDLGVRVCEKMALYDVVS 586

Qy 185 TLPQAVMGSSYGFQYSPKQRFVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVESIY 244
Db 587 TLPQAVMGSSYGFQYSPKQRFVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVESIY 646

Qy 245 QCCLDLAPARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGLVTTSCGNLTLCYLKAA 304
Db 647 QCCLDLAPARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGLVTTSCGNLTLCYLKAS 706

Qy 305 AACRAAKLQDCTMLVNGDDLWVICESAGTQBDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
Db 707 AACRAAKLQDCTMLVNGDDLWVICESAGTQBDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 766

Qy 365 ELITSCSSNSVSAHDASGKRVYLTDRPTVPLARAWEETARHTPVNSWMLGNIIMYAPTLW 424
Db 767 ELITSCSSNSVSAHDASGKRVYLTDRPTVPLARAWEETARHTPVNSWMLGNIIMYAPTLW 826

```

QY 425 ARMILWTHFFSILLAOEQLKALDCQIYGACYSIEPLDLPOIERLHGLSAPSLHSYSPG 484
DB 827 ARMILWTHFFSILLAOEQLKALDCQIYGACYSIEPLDLPOIERLHGLSAPSLHSYSPG 886
QY 485 EINRVASCLRLKGLVPLRVWRHRSVRARLLSQGAAICCKYLFNNAVRTKLTIP 544
DB 887 EINRVASCLRLKGLVPLRVWRHRSVRARLLSQGAAICCKYLFNNAVRTKLTIP 946
QY 545 AASRLDLSGFWAGYSGGDIYHLSRARPRWFMCLLLLSVGVGIYLLPNR 595
DB 947 AASRLDLSGFWAGYSGGDIYHLSRARPRWFMCLLLLSVGVGIYLLPNR 997
RESULT 6
US-08-952-981A-2
; Sequence 2, Application US/08952981A
; Patent No. 6383768
; GENERAL INFORMATION:
; APPLICANT: DE FRANCESCO, Raffaele
; APPLICANT: TOMELI, Licia
; APPLICANT: BEHRENS, Sven-Erik
; TITLE OF INVENTION: METHOD FOR REPRODUCING IN VITRO THE RNA-DEPENDENT RNA
; TITLE OF INVENTION: POLYMERASE AND TERMINAL NUCLEOTIDYL TRANSFERASE
; TITLE OF INVENTION: ACTIVITIES ENCODED BY HEPATITIS C VIRUS (HCV)
; FILE REFERENCE: IT0002P
; CURRENT APPLICATION NUMBER: US/08952,981A
; CURRENT FILING DATE: 1998-03-23
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2201
; TYPE: PRT
; ORGANISM: CDNA clone pCD (38-9.4)
US-08-952-981A-2
Query Match 95.2%; Score 3016; DB 3; Length 2201;
Best Local Similarity 96.6%; Pred. No. 1.8e-297;
Matches 571; Conservative 11; Mismatches 9; Indels 0; Gaps 0;
QY 5 SMSYTWTCALITPCAARESKLPINALSNLSLRHNLVYSTSRASISLRQKVTDFRLQVL 64
DB 1611 SMSYTWTCALITPCAARESKLPINALSNLSLRHNLVYSTSRASISLRQKVTDFRLQVL 1670
QY 65 DDHYRDLVKEMKAKASTVAKLLSVEEACKLTPPHSAKSKFGYGAADVRSLSRAVNHIR 124
DB 1671 DDHYRDLVKEMKAKASTVAKLLSVEEACKLTPPHSAKSKFGYGAADVRSLSRAVNHIR 1730
QY 125 SVWKDLEDLDTPIQTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMAIYDVVS 184
DB 1731 SVWKDLEDLDTPIQTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMAIYDVVS 1790
QY 185 TLPQAVMGSSYGFQSPQKRVBEFLVNTWKAKCPMGFSYDTRCFDSTTENDIRVEESIIY 244
DB 1791 TLPQAVMGSSYGFQSPQKRVBEFLVNTWKAKCPMGFSYDTRCFDSTTENDIRVEESIIY 1850
QY 245 OCCDLAPAROAIRSLTERLYVGGPMTNSKQNCQYRRCRAGSVLTTSCGNLTCLYKAA 304
DB 1851 OCCDLAPAROAIRSLTERLYVGGPMTNSKQNCQYRRCRAGSVLTTSCGNLTCLYKAA 1910
QY 305 AACRAKALQDCTMLVNGDDLVCICSAGTQEDAAASLRVFTTEAMTRYSAAPPDGPQPEYDL 364
DB 1911 AACRAKALQDCTMLVNGDDLVCICSAGTQEDAAASLRVFTTEAMTRYSAAPPDGPQPEYDL 1970
QY 365 ELITSCSSNVSAHDASGRVYLLTRDTVPLARAWEETARTHTPVNSWGLNIIMYAPTILW 424
DB 1971 ELITSCSSNVSAHDASGRVYLLTRDTVPLARAWEETARTHTPVNSWGLNIIMYAPTILW 2030
QY 425 ARMILWTHFFSILLAOEQLKALDCQIYGACYSIEPLDLPOIERLHGLSAPSLHSYSPG 484
DB 2031 ARMILWTHFFSILLAOEQLKALDCQIYGACYSIEPLDLPOIERLHGLSAPSLHSYSPG 2090
QY 485 EINRVASCLRLKGLVPLRVWRHRSVRARLLSQGAAICCKYLFNNAVRTKLTIP 544

DB 2091 EINRVASCLRLKGLVPLRVWRHRSVRARLLSQGAAICCKYLFNNAVRTKLTIP 2150
QY 545 AASRLDLSGFWAGYSGGDIYHLSRARPRWFMCLLLLSVGVGIYLLPNR 595
DB 2151 AASRLDLSGFWAGYSGGDIYHLSRARPRWFMCLLLLSVGVGIYLLPNR 2201
RESULT 7
US-08-324-977-36
; Sequence 36, Application US/08324977
; Patent No. 5747339
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, Mclelland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703D
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELE: 440142
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2621 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-324-977-36
Query Match 95.2%; Score 3016; DB 1; Length 2621;
Best Local Similarity 96.6%; Pred. No. 2.4e-297;

Matches 571; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

Qy 5 SMSYTWTCALITPCAAEESKLPINALSNSLLRHNLVYVSTTSRSASLRQKVKVTFDRLOVL 64
Db 2031 SMSYTWTCALITPCAAEESKLPINALSNSLLRHNLVYVSTTSRSASLRQKVKVTFDRLOVL 2090

Qy 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAADVRSLSRAVNHIR 124
Db 2091 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAADVRSLSRAVNHIR 2150

Qy 125 SVWKDLLEDTDTPITQTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 184
Db 2151 SVWKDLLEDTDTPITQTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 2210

Qy 185 TLPOAVMGSSYGFQYSPQORVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
Db 2211 TLPOAVMGSSYGFQYSPQORVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 2270

Qy 245 OCCDLAPEARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTCGNTLTCLYKAA 304
Db 2271 OCCDLAPEARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTCGNTLTCLYKAS 2330

Qy 305 AACRAAKLQDCTMLVNGDDLVCESAGTQBDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
Db 2331 AACRAAKLQDCTMLVNGDDLVCESAGTQBDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 2390

Qy 365 ELITSCSNVSVAHDASGKRVYVLTDRPTVPLARAAMETARHTPVNSWLGNIIMVAPTLW 424
Db 2391 ELITSCSNVSVAHDASGKRVYVLTDRPTVPLARAAMETARHTPVNSWLGNIIMVAPTLW 2450

Qy 425 ARMLTMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLPQIIRLHGLSFAFSLHSYSPG 484
Db 2451 ARMLTMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLPQIIRLHGLSFAFSLHSYSPG 2510

Qy 485 EIRNVASCLRLKLGVPPLRVWRHARSVRALLSQGGRAATCGKYLFWNAVTKLKLTPIP 544
Db 2511 EIRNVASCLRLKLGVPPLRVWRHARSVRALLSQGGRAATCGKYLFWNAVTKLKLTPIP 2570

Qy 545 AASRLDLSGWFVAGYSGGDIYHLSRARPRFWMCLLLLSVGVGIYLLPNR 595
Db 2571 AASRLDLSGWFVAGYSGGDIYHLSRARPRFWMCLLLLSVGVGIYLLPNR 2621

RESULT 8

US-08-384-616-36

; Sequence 36, Application US/08384616

; Patent No. 5847101

; GENERAL INFORMATION:

; APPLICANT: OKAYAMA, Hiroto

; APPLICANT: FUKU, Isao

; APPLICANT: MORI, Chisato

; APPLICANT: TAKAMIZAWA, Akahisa

; APPLICANT: YOSHIDA, Iwao

; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC

; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE

; NUMBER OF SEQUENCES: 50

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Armstrong, Westerman, Hattori, Mcleland &

; STREET: 1725 K St. N.W. Suite 1000

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20006

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/384,616

; FILING DATE:

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/769,996

; FILING DATE: 02-OCT-1991

; APPLICATION NUMBER: JP 2-167466

; FILING DATE: 25-JUN-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 2-230921

; FILING DATE: 31-AUG-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 2-305605

; FILING DATE: 09-NOV-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/635,451

; FILING DATE: 28-DEC-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: Stevens-Smith, Theresa M.

; REGISTRATION NUMBER: 36,281

; REFERENCE/DOCKET NUMBER: 900703B

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 659-2930

; TELEFAX: (202) 887-0357

; TELEX: 440142

; INFORMATION FOR SEQ ID NO: 36:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2621 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-384-616-36

Query Match 95,2%; Score 3016; DB 2; Length 2621;

Best Local Similarity 96,6%; Pred. No. 2.4e-297;

Matches 571; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

Qy 5 SMSYTWTCALITPCAAEESKLPINALSNSLLRHNLVYVSTTSRSASLRQKVKVTFDRLOVL 64

Db 2031 SMSYTWTCALITPCAAEESKLPINALSNSLLRHNLVYVSTTSRSASLRQKVKVTFDRLOVL 2090

Qy 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAADVRSLSRAVNHIR 124

Db 2091 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAADVRSLSRAVNHIR 2150

Qy 125 SVWKDLLEDTDTPITQTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 184

Db 2151 SVWKDLLEDTDTPITQTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 2210

Qy 185 TLPOAVMGSSYGFQYSPQORVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 244

Db 2211 TLPOAVMGSSYGFQYSPQORVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 2270

Qy 245 OCCDLAPEARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTCGNTLTCLYKAA 304

Db 2271 OCCDLAPEARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTCGNTLTCLYKAS 2330

Qy 305 AACRAAKLQDCTMLVNGDDLVCESAGTQBDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364

Db 2331 AACRAAKLQDCTMLVNGDDLVCESAGTQBDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 2390

Qy 365 ELITSCSNVSVAHDASGKRVYVLTDRPTVPLARAAMETARHTPVNSWLGNIIMVAPTLW 424

Db 2391 ELITSCSNVSVAHDASGKRVYVLTDRPTVPLARAAMETARHTPVNSWLGNIIMVAPTLW 2450

Qy 425 ARMLTMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLPQIIRLHGLSFAFSLHSYSPG 484

Db 2451 ARMLTMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLPQIIRLHGLSFAFSLHSYSPG 2510

Qy 485 EIRNVASCLRLKLGVPPLRVWRHARSVRALLSQGGRAATCGKYLFWNAVTKLKLTPIP 544

Db 2511 EIRNVASCLRLKLGVPPLRVWRHARSVRALLSQGGRAATCGKYLFWNAVTKLKLTPIP 2570

Qy 545 AASRLDLSGWFVAGYSGGDIYHLSRARPRFWMCLLLLSVGVGIYLLPNR 595

Db 2571 AASRLDLSGWFVAGYSGGDIYHLSRARPRFWMCLLLLSVGVGIYLLPNR 2621

RESULT 9
 US-08-904-686A-36
 ; Sequence 36, Application US/08904686A
 ; Patent No. 5998130
 ; GENERAL INFORMATION:
 ; APPLICANT: OKAYAMA, Hiroto
 ; APPLICANT: FUKU, Isao
 ; APPLICANT: MORI, Chisato
 ; APPLICANT: TAKAMIZAWA, Akahisa
 ; APPLICANT: YOSHIDA, Iwao
 ; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
 ; NUMBER OF SEQUENCES: 50
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
 ; STREET: 1725 K St. N.W. Suite 1000
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20006
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/904,686A
 ; FILING DATE: 01-AUG-1997
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/324,977
 ; FILING DATE: 18-OCT-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 2-167466
 ; FILING DATE: 25-JUN-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 2-230921
 ; FILING DATE: 31-AUG-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 2-305605
 ; FILING DATE: 09-NOV-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/099,706
 ; FILING DATE: 30-JUL-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/769,996
 ; FILING DATE: 02-OCT-1991
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/635,451
 ; FILING DATE: 28-DEC-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McLeland, Le-Nhung
 ; REGISTRATION NUMBER: 31,541
 ; REFERENCE/DOCKET NUMBER: 900703G
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 659-2930
 ; TELEFAX: (202) 887-0357
 ; INFORMATION FOR SEQ ID NO: 36:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2621 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-08-904-686A-36

Query Match 95.2%; Score 3016; DB 2; Length 2621;
 Best Local Similarity 96.6%; Pred. No. 2.4e-297;
 Matches 571; Conservative 11; Mismatches 9; Indels 0; Gaps 0;
 5 SMSYTWGALITPCAAESKLPINALNSLLRHHNMVYATTSRSASLRQKVTDFRLQVL 64

Db 2031 SMSYTWGALITPCAAESKLPINALNSLLRHHNMVYATTSRSASLRQKVTDFRLQVL 2090
 QY 65 DDHYRDLVKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAADVRSLSRAVNHIR 124
 Db 2091 DDHYRDLVKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAADVRSLSRAVNHIR 2150
 QY 125 SVWKOLLEDTPTPIOTTIMAKNEVFCVQPEKGRKPARLIVFPDGLGVRVCBKWALYDVVS 184
 Db 2151 SVWKOLLEDTPTPIOTTIMAKNEVFCVQPEKGRKPARLIVFPDGLGVRVCBKWALYDVVS 2210
 QY 185 TLPQAVMGSSYGYFQYSPQGRVFEFLVNTWKAKCKPMGFSDTCEFDSTVTENDIRVEESIY 244
 Db 2211 TLPQAVMGSSYGYFQYSPQGRVFEFLVNTWKAKCKPMGFSDTCEFDSTVTENDIRVEESIY 2270
 QY 245 QCCDLAPARQAIRSLTERLYVGGPMNTSKGQNCYRRCRAGSVLTTCGNTLTCTYKAA 304
 Db 2271 QCCDLAPARQAIRSLTERLYVGGPMNTSKGQNCYRRCRAGSVLTTCGNTLTCTYKAA 2330
 QY 305 AACRAAKLQDCTMLVNGDVLVVICESAGTQEDAAASLRVFTTEAMTRYSPAPGPPQPEYDL 364
 Db 2331 AACRAAKLQDCTMLVNGDVLVVICESAGTQEDAAASLRVFTTEAMTRYSPAPGPPQPEYDL 2390
 QY 365 ELITSCSSNVSAHDASGRVYLTTRDPTVPLARAATETARTHTPVNSWLGNIIMYAPTILW 424
 Db 2391 ELITSCSSNVSAHDASGRVYLTTRDPTVPLARAATETARTHTPVNSWLGNIIMYAPTILW 2450
 QY 425 ARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIRLHGLSASFSLHSYSPG 484
 Db 2451 ARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIRLHGLSASFSLHSYSPG 2510
 QY 485 EIRNVASCLRLKGLVPPPLRVWRHARSVRALKLSQGGRAAICGKYLFWNAVRTKLTLTPIP 544
 Db 2511 EIRNVASCLRLKGLVPPPLRVWRHARSVRALKLSQGGRAAICGKYLFWNAVRTKLTLTPIP 2570
 QY 545 AASRLDLSGWFVAGYSGGDIYHSLSRAPRPFMLCLLLSVGVGYLLPNR 595
 Db 2571 AASRLDLSGWFVAGYSGGDIYHSLSRAPRPFMLCLLLSVGVGYLLPNR 2621

RESULT 10
 US-09-315-850-36
 ; Sequence 36, Application US/09315850
 ; Patent No. 6217872
 ; GENERAL INFORMATION:
 ; APPLICANT: OKAYAMA, Hiroto
 ; APPLICANT: FUKU, Isao
 ; APPLICANT: MORI, Chisato
 ; APPLICANT: TAKAMIZAWA, Akahisa
 ; APPLICANT: YOSHIDA, Iwao
 ; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
 ; NUMBER OF SEQUENCES: 50
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
 ; STREET: 1725 K St. N.W. Suite 1000
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20006
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/315,850
 ; FILING DATE:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/904,686
 ; FILING DATE: 01-AUG-1997
 ; APPLICATION NUMBER: US 08/324,977
 ; FILING DATE: 18-OCT-1994

; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 2-167466
 ; FILING DATE: 25-JUN-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 2-230921
 ; FILING DATE: 31-AUG-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 2-305605
 ; FILING DATE: 09-NOV-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/099,706
 ; FILING DATE: 30-JUL-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/769,996
 ; FILING DATE: 02-OCT-1991
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/635,451
 ; FILING DATE: 28-DEC-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McLeLand, Le-Nhung
 ; REGISTRATION NUMBER: 31,541
 ; REFERENCE/DOCKET NUMBER: 900703G
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 659-2930
 ; TELEFAX: (202) 887-0357
 ; INFORMATION FOR SEQ ID NO: 36:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2621 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-315-850-36

Query Match 95.2%; Score 3016; DB 3; Length 2621;
 Best Local Similarity 96.6%; Pred. No. 2.4e-297;
 Matches 571; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

QY 5 SMSYTWTCALITPCAABESKLPINALSNLHHNLVYTTSSRSASLRQKKVTFDRLOVL 64
 DB 2031 SMSYTWTCALITPCAABESKLPINALSNLHHNLVYTTSSRSASLRQKKVTFDRLOVL 2090
 QY 65 DDHYRDVLKEMKAKASTYKALLSVEEACKLTPPHSAKSKFGYGAQVRSLSRAVNHIR 124
 DB 2091 DDHYRDVLKEMKAKASTYKALLSVEEACKLTPPHSAKSKFGYGAQVRSLSRAVNHIR 2150
 QY 125 SWKDLLEDDTPIOTTIMAKNEVFCVQPEKGRKPARLIIVPDLGVRVCEKMALYDVVS 184
 DB 2151 SWKDLLEDDTPIOTTIMAKNEVFCVQPEKGRKPARLIIVPDLGVRVCEKMALYDVVS 2210
 QY 185 TLPOAVMGSSYGFQYSPKORVEFLVNTWKAKCPMGFSYDTRCPDSTVTENDIRVEESIY 244
 DB 2211 TLPOAVMGSSYGFQYSPKORVEFLVNTWKAKCPMGFSYDTRCPDSTVTENDIRVEESIY 2270
 QY 245 QCDDLAPEARQAIRSLTERLYVGGPMTNSKQNCYRRCRASGVLTISCGNTLTICLKAA 304
 DB 2271 QCDDLAPEARQAIRSLTERLYVGGPMTNSKQNCYRRCRASGVLTISCGNTLTICLKAA 2330
 QY 305 AACRAAKLQDCTMLVNGDDLVCSSAGTQEDAAASLRVFTAMTRYSAAPPDPPQPEYDL 364
 DB 2331 AACRAAKLQDCTMLVNGDDLVCSSAGTQEDAAASLRVFTAMTRYSAAPPDPPQPEYDL 2390
 QY 365 ELITSCSNVSAVHDASGRVYLLTRDTPVLARAWEATARTHPVNSWLNIIIVAPTILW 424
 DB 2391 ELITSCSNVSAVHDASGRVYLLTRDTPVLARAWEATARTHPVNSWLNIIIVAPTILW 2450
 QY 425 ARMILMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPOIERLHGLSAFSLHSYSPG 484
 DB 2451 ARMILMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPOIERLHGLSAFSLHSYSPG 2510
 QY 485 EINRVASCLRLKGVPLPVRHRRARSVRAKLLSQGGRAAICCKYLFNVAWTKLKLTPIP 544
 DB 2511 EINRVASCLRLKGVPLPVRHRRARSVRAKLLSQGGRAAICCKYLFNVAWTKLKLTPIP 2570

QY 545 AASRLDLSGWFVAGYSGGDIYHLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
 DB 2571 AASRLDLSGWFVAGYSGGDIYHLSRARPRWFMLCLLLLSVGVGIYLLPNR 2621
 RESULT 11
 US-08-324-977-2
 ; Sequence 2, Application US/08324977
 ; Patent No. 5747339
 ; GENERAL INFORMATION:
 ; APPLICANT: OKAYAMA, Hiroto
 ; APPLICANT: FUKU, Isao
 ; APPLICANT: MORI, Chisato
 ; APPLICANT: TAKAMIZAWA, Akahisa
 ; APPLICANT: YOSHIDA, Iwao
 ; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
 ; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
 ; NUMBER OF SEQUENCES: 50
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLand &
 ; ADDRESSEE: Naughton
 ; STREET: 1725 K St. N.W. Suite 1000
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20006
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/324,977
 ; FILING DATE: 18-OCT-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 2-167466
 ; FILING DATE: 25-JUN-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 2-230921
 ; FILING DATE: 31-AUG-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 2-305605
 ; FILING DATE: 09-NOV-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/099,706
 ; FILING DATE: 30-JUL-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/769,996
 ; FILING DATE: 02-OCT-1991
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/635,451
 ; FILING DATE: 28-DEC-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Stevens-Smith, Theresa M.
 ; REGISTRATION NUMBER: 36,281
 ; REFERENCE/DOCKET NUMBER: 900703D
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 659-2930
 ; TELEFAX: (202) 887-0357
 ; TELEX: 440142
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3010 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-324-977-2

Query Match 95.2%; Score 3016; DB 1; Length 3010;
 Best Local Similarity 96.6%; Pred. No. 3e-297;
 Matches 571; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

QY 5 SMSYTWTCALITPCAABESKLPINALSNLHHNLVYTTSSRSASLRQKKVTFDRLOVL 64

Db 2420 SMSYTWGALITPCAAEESKLPINALSNLSLRHNNVYATTSRAGLRQKQKTFORLQVL 2479
QY 65 DDHYRDVLKEMKAKASTVKALLSVEEACKLTPPHSAKSKFGYGAADVRSLSRAVNHIR 124
Db 2480 DDHYRDVLKEMKAKASTVKALLSVEEACKLTPPHSAKSKFGYGAADVNLSSKAVNHIIH 2539
QY 125 SVWKDLLEDTPPIQTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKALYDVVYS 184
Db 2540 SVWKDLLEDTPPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKALYDVVYS 2599
QY 185 TLPOAVMGSSYGFQYSPQORVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
Db 2600 TLPOAVMGSSYGFQYSPQORVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 2659
QY 245 OCCDLAPARQAIRSLTERLYVGGPMTNSKGQNCYRRCRAGSVLTTCGNTLTCTYLKAA 304
Db 2660 OCCDLAPARQAIRSLTERLYVGGPMTNSKGQNCYRRCRAGSVLTTCGNTLTCTYLKAS 2719
QY 305 AACRAAKLQDCTMLVNGDDLVCICESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
Db 2720 AACRAAKLQDCTMLVNGDDLVCICESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 2779
QY 365 ELITSCSSNVSAHDASGKRVYLTREDPTVPLARAAMETARHTPVNSWLGNIIMYAPTILW 424
Db 2780 ELITSCSSNVSAHDASGKRVYLTREDPTVPLARAAMETARHTPVNSWLGNIIMYAPTILW 2839
QY 425 ARMLIMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIERLHGLSAPSLHSYSPG 484
Db 2840 ARMLIMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIERLHGLSAPSLHSYSPG 2899
QY 485 EIRNVASCLRLKGLVPPPLRVWRHARSVRALKLSQGGRAAICGKYLFWNAVRTKLTPTIP 544
Db 2900 EIRNVASCLRLKGLVPPPLRVWRHARSVRALKLSQGGRAATCGKYLFWNAVRTKLTPTIP 2959
QY 545 AASRLDLSGWFVAGYSGGDIYHLSRARPRWFMCLCLLLSVGVGIYLLPNR 595
Db 2960 AASRLDLSGWFVAGYSGGDIYHLSRARPRWFMCLCLLLSVGVGIYLLPNR 3010

RESULT 12

US-08-324-977-14
; Sequence 14, Application US/08324977
; Patent No. 5747339
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeand &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3010 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-324-977-14

Query Match 95.2%; Score 3016; DB 1; Length 3010;
Best Local Similarity 96.6%; Pred. No. 3e-297;
Matches 571; Conservative 11; Mismatches 9; Indels 0; Gaps 0;
QY 5 SMSYTWGALITPCAAEESKLPINALSNLSLRHNNVYATTSRAGLRQKQKTFORLQVL 64
Db 2420 SMSYTWGALITPCAAEESKLPINALSNLSLRHNNVYATTSRAGLRQKQKTFORLQVL 2479
QY 65 DDHYRDVLKEMKAKASTVKALLSVEEACKLTPPHSAKSKFGYGAADVRSLSRAVNHIR 124
Db 2480 DDHYRDVLKEMKAKASTVKALLSVEEACKLTPPHSAKSKFGYGAADVNLSSKAVNHIIH 2539
QY 125 SVWKDLLEDTPPIQTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKALYDVVYS 184
Db 2540 SVWKDLLEDTPPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKALYDVVYS 2599
QY 185 TLPOAVMGSSYGFQYSPQORVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
Db 2600 TLPOAVMGSSYGFQYSPQORVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 2659
QY 245 OCCDLAPARQAIRSLTERLYVGGPMTNSKGQNCYRRCRAGSVLTTCGNTLTCTYLKAA 304
Db 2660 OCCDLAPARQAIRSLTERLYVGGPMTNSKGQNCYRRCRAGSVLTTCGNTLTCTYLKAS 2719
QY 305 AACRAAKLQDCTMLVNGDDLVCICESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
Db 2720 AACRAAKLQDCTMLVNGDDLVCICESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 2779
QY 365 ELITSCSSNVSAHDASGKRVYLTREDPTVPLARAAMETARHTPVNSWLGNIIMYAPTILW 424
Db 2780 ELITSCSSNVSAHDASGKRVYLTREDPTVPLARAAMETARHTPVNSWLGNIIMYAPTILW 2839
QY 425 ARMLIMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIERLHGLSAPSLHSYSPG 484
Db 2840 ARMLIMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIERLHGLSAPSLHSYSPG 2899
QY 485 EIRNVASCLRLKGLVPPPLRVWRHARSVRALKLSQGGRAAICGKYLFWNAVRTKLTPTIP 544
Db 2900 EIRNVASCLRLKGLVPPPLRVWRHARSVRALKLSQGGRAATCGKYLFWNAVRTKLTPTIP 2959
QY 545 AASRLDLSGWFVAGYSGGDIYHLSRARPRWFMCLCLLLSVGVGIYLLPNR 595
Db 2960 AASRLDLSGWFVAGYSGGDIYHLSRARPRWFMCLCLLLSVGVGIYLLPNR 3010

RESULT 13
US-08-384-616-2
; Sequence 2, Application US/08384616
; Patent No. 5847101
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/384,616
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3010 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-384-616-2
Query Match 95.2%; Score 3016; DB 2; Length 3010;
Best Local Similarity 96.6%; Pred. No. 3e-297;
Matches 571; Conservative 11; Mismatches 9; Indels 0; Gaps 0;
QY 5 SMSYTWTCALITPCAARESKLPINALSNLLRHHNLVYTSRSASLRQKKVTFDRLOVL 64
DB 2420 SMSYTWTCALITPCAARESKLPINALSNLLRHHNLVYTSRSASLRQKKVTFDRLOVL 2479
QY 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
DB 2480 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSKAVNH 2539

125 SVWKDLLEDTDTPITQTTIMAKNEVFCVQPEKGGKRPARLIVFPDLGVRVCEKMALYDVVS 184
DB 2540 SVWKDLLEDVTPTDITTIMAKNEVFCVQPEKGGKRPARLIVFPDLGVRVCEKMALYDVVS 2599
QY 185 TLPQAVMGSSYGFQYSPKQRFVFLVNTWKAKCKPMGFSYDTRCFDSTVTENDIRVESIY 244
DB 2600 TLPQAVMGSSYGFQYSPQRFVFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRVESIY 2659
QY 245 QCCDLAPEARQAIKSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTSCTGNTLTCYLKAA 304
DB 2660 QCCDLAPEARQAIKSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTSCTGNTLTCYLKAS 2719
QY 305 AACRAAKLQDCTMLVNGDDLVVICESAGTQDAASLVFTEAMTRYSAAPPDPPQPEYDL 364
DB 2720 AACRAAKLQDCTMLVNGDDLVVICESAGTQDAASLVFTEAMTRYSAAPPDPPQPEYDL 2779
QY 365 ELITSCSSNVSAHDSAGKRVYVLTTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTLW 424
DB 2780 ELITSCSSNVSAHDSAGKRVYVLTTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTLW 2839
QY 425 ARMILMTHFFSILLAQOLEKALDCQIYGACYSIEPLDLPQIIERLHGLSFAFSLHSYSPG 484
DB 2840 ARMILMTHFFSILLAQOLEKALDCQIYGACYSIEPLDLPQIIERLHGLSFAFSLHSYSPG 2899
QY 485 EINEVASCLRKLGVPPLRVVHRARSVRKLLSOGGAAICGKYLENNAVETKLTLP 544
DB 2900 EINEVASCLRKLGVPPLRVVHRARSVRKLLSOGGAAICGKYLENNAVETKLTLP 2959
QY 545 AASRLDLSGWFVAGYSGGDIYHLSLRARPRFWMCLLLLSVGVGIYLLPNR 595
DB 2960 AASRLDLSGWFVAGYSGGDIYHLSLRARPRFWMCLLLLSVGVGIYLLPNR 3010

RESULT 14
US-08-384-616-14
; Sequence 14, Application US/08384616
; Patent No. 5847101
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/384,616
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3010 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-384-616-2

; APPLICATION NUMBER: JP 2-305605
 ; FILING DATE: 09-NOV-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/635,451
 ; FILING DATE: 28-DEC-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Stevens-Smith, Theresa M.
 ; REGISTRATION NUMBER: 36,281
 ; REFERENCE/DOCKET NUMBER: 900703B
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 659-2930
 ; TELEFAX: (202) 887-0357
 ; TELEX: 440142
 ; INFORMATION FOR SEQ ID NO: 14:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3010 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-384-616-14

Query Match 95.2%; Score 3016; DB 2; Length 3010;
 Best Local Similarity 96.6%; Pred. No. 3e-297;
 Matches 571; Conservative 11; Mismatches 9; Indels 0; Gaps 0;
 QY 5 SMSYTWTCALITPCAABESKLPINALSNLLRHHNLVYTTSSASLRQKKVTFDLQVL 64
 Db 2420 SMSYTWTCALITPCAABESKLPINALSNLLRHHNMVYATTSSAGLRQKKVTFDLQVL 2479
 QY 65 DDHYRDVLKEMKAKASTVKALLSVEEACKLTPPHSAKSKFGYGAQVRSLSRAVNHIR 124
 Db 2480 DDHYRDVLKEMKAKASTVKALLSVEEACKLTPPHSAKSKFGYGAQVRSLSRAVNHIR 2539
 QY 125 SVWKDLLEDTPITOTTIMAKNEVFCVQPEKGRKPARLI VFPDLGVRCCKMALYDVVS 184
 Db 2540 SVWKDLLEDTPITOTTIMAKNEVFCVQPEKGRKPARLI VFPDLGVRCCKMALYDVVS 2599
 QY 185 TLPAQVMSSYGFQYSPQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
 Db 2600 TLPAQVMSSYGFQYSPQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 2659
 QY 245 QCCDLAPARQAIRSLTRLVYGGPMWNSKQNGCYRCRAGVLTTSCGNTLTCLYKAA 304
 Db 2660 QCCDLAPARQAIRSLTRLVYGGPMWNSKQNGCYRCRAGVLTTSCGNTLTCLYKAS 2719
 QY 305 AACRAAKLQDCTMLVNGDVLVVCESAGTOEDASLRVFTAMTRYSAAPGDPPOPEYDL 364
 Db 2720 AACRAAKLQDCTMLVNGDVLVVCESAGTOEDASLRVFTAMTRYSAAPGDPPOPEYDL 2779
 QY 365 ELITSCSNVSVAHDAKGRVYLTTRDPTVPLARAAMETARHTPVNSWLGNIIMVAPTLW 424
 Db 2780 ELITSCSNVSVAHDAKGRVYLTTRDPTVPLARAAMETARHTPVNSWLGNIIMVAPTLW 2839
 QY 425 ARMLTMTHFFSILLAEQLEKALDQIYGACYSIEPLDLPQIERLHGLSAFSLHSYSPG 484
 Db 2840 ARMLTMTHFFSILLAEQLEKALDQIYGACYSIEPLDLPQIERLHGLSAFSLHSYSPG 2899
 QY 485 EINRVASCLRLKGVPPPLRVWRHARSVRKALLSOGGAAICGKYLFWAVRTKLKLTIP 544
 Db 2900 EINRVASCLRLKGVPPPLRVWRHARSVRKALLSOGGAAICGKYLFWAVRTKLKLTIP 2959
 QY 545 AASRLDLSGFWAGVSGGDIYHLSRARPFRWMLCLLLSVGVGYLLLPNR 595
 Db 2960 AASRLDLSGFWAGVSGGDIYHLSRARPFRWMLCLLLSVGVGYLLLPNR 3010

RESULT 15
 US-08-904-686A-2
 ; Sequence 2, Application US/08904686A
 ; Patent No. 5998130
 ; GENERAL INFORMATION:
 ; APPLICANT: OKAYAMA, Hiroto
 ; APPLICANT: FUKU, Isao

; APPLICANT: MORI, Chisato
 ; APPLICANT: TAKAMIZAWA, Akahisa
 ; APPLICANT: YOSHIDA, Iwao
 ; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
 ; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
 ; NUMBER OF SEQUENCES: 50
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Armstrong, Westerman, Hattori, McLealand &
 ; ADDRESSEE: Naughton
 ; STREET: 1725 K St. N.W. Suite 1000
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20006
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/904,686A
 ; FILING DATE: 01-AUG-1997
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/324,977
 ; FILING DATE: 18-OCT-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 2-167466
 ; FILING DATE: 25-JUN-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 2-230921
 ; FILING DATE: 31-AUG-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 2-305605
 ; FILING DATE: 09-NOV-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/099,706
 ; FILING DATE: 30-JUL-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/769,996
 ; FILING DATE: 02-OCT-1991
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/635,451
 ; FILING DATE: 28-DEC-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McLealand, Le-Nhung
 ; REGISTRATION NUMBER: 31,541
 ; REFERENCE/DOCKET NUMBER: 900703G
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 659-2930
 ; TELEFAX: (202) 887-0357
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3010 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-904-686A-2

Query Match 95.2%; Score 3016; DB 2; Length 3010;
 Best Local Similarity 96.6%; Pred. No. 3e-297;
 Matches 571; Conservative 11; Mismatches 9; Indels 0; Gaps 0;
 QY 5 SMSYTWTCALITPCAABESKLPINALSNLLRHHNLVYTTSSASLRQKKVTFDLQVL 64
 Db 2420 SMSYTWTCALITPCAABESKLPINALSNLLRHHNMVYATTSSAGLRQKKVTFDLQVL 2479
 QY 65 DDHYRDVLKEMKAKASTVKALLSVEEACKLTPPHSAKSKFGYGAQVRSLSRAVNHIR 124
 Db 2480 DDHYRDVLKEMKAKASTVKALLSVEEACKLTPPHSAKSKFGYGAQVRSLSRAVNHIR 2539
 QY 125 SVWKDLLEDTPITOTTIMAKNEVFCVQPEKGRKPARLI VFPDLGVRCCKMALYDVVS 184
 Db 2540 SVWKDLLEDTPITOTTIMAKNEVFCVQPEKGRKPARLI VFPDLGVRCCKMALYDVVS 2599

Qy	185	TLPOAVMGSSYGFQYSPKORVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY	244
Db	2600	TLPOVVMGSSYGFQYSPQORVEFLVNTWKSKNPMGFSYDTRCFDSTVTENDIRVEESIY	2659
Qy	245	QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNGCYRRCRASGVLTTSCGNTLTCYLKAA	304
Db	2660	QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNGCYRRCRASGVLTTSCGNTLTCYLKAS	2719
Qy	305	AACRAAKLQDCTMLVNGDDLVI CESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL	364
Db	2720	AACRAAKLQDCTMLVNGDDLVI CESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL	2779
Qy	365	ELITSCSSNVSAHDASGKRVYVLTTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTLW	424
Db	2780	ELITSCSSNVSAHDASGKRVYVLTTRDPTTPLARAAMETARHTPVNSWLGNIIMYAPTLW	2839
Qy	425	ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPOI IERLHGLSAFSLHSYSPG	484
Db	2840	ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPOI IERLHGLSAFSLHSYSPG	2899
Qy	485	EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLEFNNAVTKLKLTPIP	544
Db	2900	EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLEFNNAVTKLKLTPIP	2959
Qy	545	AASRLDLSGWFVAGYSGGDIYHSLSRAPRWFMLCLLLLSVGVGIYLLPNR	595
Db	2960	AASRLDLSGWFVAGYSGGDIYHSLSRAPRWFMLCLLLLSVGVGIYLLPNR	3010

Search completed: September 22, 2005, 14:47:44
Job time : 30 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2005, 14:43:14 ; Search time 76 Seconds
(without alignments)

3218.609 Million cell updates/sec

Title: US-10-712-479-2

Perfect score: 3167

Sequence: 1 MASMSMTWTGALITPCAA.....LLSVGVGIYLLPNRHHHHH 601

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1826521 seqs, 407012169 residues

Total number of hits satisfying chosen parameters: 1826521

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3167	100.0	601	16	US-10-712-479-2
2	3034.5	95.8	580	16	US-10-712-479-4
3	3016	95.2	591	13	US-10-085-476-1
4	3016	95.2	2201	13	US-10-085-476-2
5	3014	95.2	593	14	US-10-241-872-23
6	3012	95.1	1985	16	US-10-492-178-6
7	3011	95.1	591	16	US-10-471-164-5
8	3011	95.1	621	9	US-09-898-297-1
9	3011	95.1	621	14	US-10-198-680A-1
10	3011	95.1	621	15	US-10-238-282-1
11	3011	95.1	621	15	US-10-198-259A-1
					Sequence 2, Appli
					Sequence 4, Appli
					Sequence 1, Appli
					Sequence 2, Appli
					Sequence 23, Appli
					Sequence 6, Appli
					Sequence 5, Appli
					Sequence 1, Appli
					Sequence 1, Appli
					Sequence 1, Appli

12	3011	95.1	621	15	US-10-198-384A-1	Sequence 1, Appli
13	3011	95.1	621	16	US-10-471-164-1	Sequence 1, Appli
14	3011	95.1	621	16	US-10-851-710-1	Sequence 1, Appli
15	3011	95.1	627	9	US-09-838-386-12	Sequence 12, Appli
16	3011	95.1	627	16	US-10-471-164-6	Sequence 6, Appli
17	3006	94.9	2865	9	US-09-742-659-6	Sequence 6, Appli
18	3005	94.9	1985	14	US-10-259-275-42	Sequence 42, Appli
19	3005	94.9	1985	16	US-10-639-150-2	Sequence 2, Appli
20	3005	94.9	1985	20	US-11-006-313-42	Sequence 42, Appli
21	3005	94.9	2201	13	US-10-029-907-3	Sequence 3, Appli
22	3005	94.9	2201	14	US-10-309-561-3	Sequence 3, Appli
23	3005	94.9	2201	16	US-10-789-355-3	Sequence 3, Appli
24	3005	94.9	2201	16	US-10-686-835-3	Sequence 3, Appli
25	3005	94.9	3010	15	US-10-467-000-1	Sequence 1, Appli
26	2391	94.4	1985	16	US-10-492-178-1	Sequence 1, Appli
27	2987	94.3	593	14	US-10-241-872-12	Sequence 12, Appli
28	2978	94.0	2985	14	US-10-259-275-40	Sequence 40, Appli
29	2978	94.0	2985	20	US-11-006-313-40	Sequence 40, Appli
30	2972	93.8	3010	16	US-10-333-449A-34	Sequence 34, Appli
31	2937.5	92.8	576	17	US-10-842-046-1	Sequence 1, Appli
32	2927	92.4	590	10	US-09-194-949-15	Sequence 15, Appli
33	2927	92.4	590	17	US-10-664-391-15	Sequence 15, Appli
34	2925.5	92.4	578	14	US-10-211-455-1	Sequence 1, Appli
35	2925.5	92.4	578	16	US-10-471-164-3	Sequence 3, Appli
36	2904	91.7	591	14	US-10-211-455-2	Sequence 2, Appli
37	2904	91.7	591	14	US-10-211-455-4	Sequence 4, Appli
38	2904	91.7	591	16	US-10-471-164-2	Sequence 2, Appli
39	2879	90.9	576	14	US-10-170-131-1	Sequence 1, Appli
40	2786	88.0	591	9	US-09-733-183A-2	Sequence 2, Appli
41	2786	88.0	591	14	US-10-342-372-2	Sequence 2, Appli
42	2776	87.7	3011	15	US-10-296-734-406	Sequence 406, App
43	2774	87.6	3011	9	US-09-742-659-4	Sequence 4, Appli
44	2774	87.6	3011	9	US-09-952-572-9	Sequence 9, Appli
45	2774	87.6	3011	9	US-09-747-419-20	Sequence 20, Appli

ALIGNMENTS

RESULT 1

US-10-712-479-2
; Sequence 2, Application US/10712479
; Publication No. US20040209283A1
; GENERAL INFORMATION:
; APPLICANT: Yagi, Yoshihiko
; APPLICANT: Sheets, Michael P.
; APPLICANT: Wells, Peter A.
; APPLICANT: Shelly, John A.
; APPLICANT: Poorman, Roger A.
; APPLICANT: Epps, Dennis E.
; TITLE OF INVENTION: A Continuous-Read Assay for the Detection of De Novo HCV RNA Polym
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 02-333-A
; CURRENT APPLICATION NUMBER: US/10/712.479
; CURRENT FILING DATE: 2003-11-13
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 601
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-712-479-2

Query Match	100.0%	Score 3167;	DB 16;	Length 601;
Best Local Similarity	100.0%	Pred. No. 4e-280;		
Matches	601;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Qy	1	MASMSMTWTGALITPCAAESKLPINALNSLRHNLVYSTRSGASLRQKVTEDR	60	
Db	1	MASMSMTWTGALITPCAAESKLPINALNSLRHNLVYSTRSGASLRQKVTEDR	60	
Qy	61	LQVLDDHYRDLKEMKAKASTVKALLSVERACKLTPPHSAKSGFYGAKDVRSLSRAV	120	

```
Db 61 LOVLDDHYRDLVKEMKAKASTVKAALLSVEBACKLTPPHSAKSKFGYGAKDVRSLSSRAV 120
Qy 121 NHIRSVKDLLEDTDTPITQTTIMAKNEVFCVQPEKGGKPARLLIVFPDLGVRVCEKNALY 180
Db 121 NHIRSVKDLLEDTDTPITQTTIMAKNEVFCVQPEKGGKPARLLIVFPDLGVRVCEKNALY 180
Qy 181 DWSTLPAQVNGSSYGFOYSPKQRFVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVE 240
Db 181 DWSTLPAQVNGSSYGFOYSPKQRFVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVE 240
Qy 241 ESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNGYRRCRASGVLTTCGNTLTCY 300
Db 241 ESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNGYRRCRASGVLTTCGNTLTCY 300
Qy 301 LKAAAACRAAKLODCTMLVNGDDLVIWICESAGTQEDAAASLRFVTEAMTRYSAAPGDPPOP 360
Db 301 LKAAAACRAAKLODCTMLVNGDDLVIWICESAGTQEDAAASLRFVTEAMTRYSAAPGDPPOP 360
Qy 361 EYDLELITSCSSNVSAHDASGKRVYVLTTRDPTVPLARAAMETARHTFPVNSWLGNIIMYA 420
Db 361 EYDLELITSCSSNVSAHDASGKRVYVLTTRDPTVPLARAAMETARHTFPVNSWLGNIIMYA 420
Qy 421 PTLWARMILMTHFFSILLAOBLEKALDCQIYGACYSIEPLDLPOIIRLHGLSAFSLHS 480
Db 421 PTLWARMILMTHFFSILLAOBLEKALDCQIYGACYSIEPLDLPOIIRLHGLSAFSLHS 480
Qy 481 YSPGEINRVASCLRLKGLVPPPLRVHRARSVRKALLSQGGRAAICGKYLFWNAVRTKLL 540
Db 481 YSPGEINRVASCLRLKGLVPPPLRVHRARSVRKALLSQGGRAAICGKYLFWNAVRTKLL 540
Qy 541 TPIPAASRLDLSGMFVAGYSGGDIYHLSLRARPRWFMCLLLLSVGVGIYLLPNRHHHH 600
Db 541 TPIPAASRLDLSGMFVAGYSGGDIYHLSLRARPRWFMCLLLLSVGVGIYLLPNRHHHH 600
Qy 601 H 601
Db 601 H 601

RESULT 2
US-10-712-479-4
; Sequence 4, Application US/10712479
; Publication No. US20040209283A1
; GENERAL INFORMATION:
; APPLICANT: Yagi, Yoshihiko
; APPLICANT: Sheets, Michael P.
; APPLICANT: Wells, Peter A.
; APPLICANT: Shelly, John A.
; APPLICANT: Poorman, Roger A.
; APPLICANT: Edge, Dennis E.
; TITLE OF INVENTION: A Continuous-Read Assay for the Detection of De Novo HCV RNA Poly
; FILE REFERENCE: 02-333-A
; CURRENT APPLICATION NUMBER: US/10/712,479
; CURRENT FILING DATE: 2003-11-13
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 580
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: C-terminally truncated HCV NS5B polymerase (C delta 21 NS5B)
US-10-712-479-4
```

```
Query Match 95.8%; Score 3034.5; DB 16; Length 580;
Best Local Similarity 96.5%; Pred. No. 4.9e-268;
Matches 580; Conservative 0; Mismatches 0; Indels 21; Gaps 1;

Qy 1 MASMSMTWTGALITPCAABESKLPINALNSLRRHNLVYSTTSRASLRQKKVTPDR 60
Db 1 MASMSMTWTGALITPCAABESKLPINALNSLRRHNLVYSTTSRASLRQKKVTPDR 60
```

```
Qy 61 LOVLDDHYRDLVKEMKAKASTVKAALLSVEBACKLTPPHSAKSKFGYGAKDVRSLSSRAV 120
Db 61 LOVLDDHYRDLVKEMKAKASTVKAALLSVEBACKLTPPHSAKSKFGYGAKDVRSLSSRAV 120
Qy 121 NHIRSVKDLLEDTDTPITQTTIMAKNEVFCVQPEKGGKPARLLIVFPDLGVRVCEKNALY 180
Db 121 NHIRSVKDLLEDTDTPITQTTIMAKNEVFCVQPEKGGKPARLLIVFPDLGVRVCEKNALY 180
Qy 181 DWSTLPAQVNGSSYGFOYSPKQRFVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVE 240
Db 181 DWSTLPAQVNGSSYGFOYSPKQRFVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVE 240
Qy 241 ESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNGYRRCRASGVLTTCGNTLTCY 300
Db 241 ESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNGYRRCRASGVLTTCGNTLTCY 300
Qy 301 LKAAAACRAAKLODCTMLVNGDDLVIWICESAGTQEDAAASLRFVTEAMTRYSAAPGDPPOP 360
Db 301 LKAAAACRAAKLODCTMLVNGDDLVIWICESAGTQEDAAASLRFVTEAMTRYSAAPGDPPOP 360
Qy 361 EYDLELITSCSSNVSAHDASGKRVYVLTTRDPTVPLARAAMETARHTFPVNSWLGNIIMYA 420
Db 361 EYDLELITSCSSNVSAHDASGKRVYVLTTRDPTVPLARAAMETARHTFPVNSWLGNIIMYA 420
Qy 421 PTLWARMILMTHFFSILLAOBLEKALDCQIYGACYSIEPLDLPOIIRLHGLSAFSLHS 480
Db 421 PTLWARMILMTHFFSILLAOBLEKALDCQIYGACYSIEPLDLPOIIRLHGLSAFSLHS 480
Qy 481 YSPGEINRVASCLRLKGLVPPPLRVHRARSVRKALLSQGGRAAICGKYLFWNAVRTKLL 540
Db 481 YSPGEINRVASCLRLKGLVPPPLRVHRARSVRKALLSQGGRAAICGKYLFWNAVRTKLL 540
Qy 541 TPIPAASRLDLSGMFVAGYSGGDIYHLSLRARPRWFMCLLLLSVGVGIYLLPNRHHHH 600
Db 541 TPIPAASRLDLSGMFVAGYSGGDIYHLSLRARPRWFMCLLLLSVGVGIYLLPNRHHHH 600
Qy 601 H 601
Db 580 H 580
```

```
RESULT 3
US-10-085-476-1
; Sequence 1, Application US/10085476
; Publication No. US20020164722A1
; GENERAL INFORMATION:
; APPLICANT: De Francesco, Raffaele
; APPLICANT: Tomei, Licia
; APPLICANT: Behrens, Sven-Erik
; TITLE OF INVENTION: METHOD FOR REPRODUCING IN VITRO THE
; TITLE OF INVENTION: RNA-DEPENDENT RNA POLYMERASE AND TERMINAL NUCLEOTIDYL
; FILE REFERENCE: IT0002PCA
; CURRENT APPLICATION NUMBER: US/10/085,476
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: 08/952,981
; PRIOR FILING DATE: 1998-03-23
; PRIOR APPLICATION NUMBER: PCT/IT96/00106
; PRIOR FILING DATE: 1996-05-24
; PRIOR APPLICATION NUMBER: RM95A000343
; PRIOR FILING DATE: 1995-05-25
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-085-476-1
```

```
Query Match 95.2%; Score 3016; DB 13; Length 591;
Best Local Similarity 96.6%; Pred. No. 2.5e-266;
Matches 571; Conservative 11; Mismatches 9; Indels 0; Gaps 0;
```

```
QY 5 SMSYTWTCALITPCAABESKLPINALSNSLLRHNLVYTTSSRSASLRQKKVTFDRLOVL 64
Db 1 SMSYTWTCALITPCAABESKLPINALSNSLLRHNLVYTTSSRSASLRQKKVTFDRLOVL 60
QY 65 DDHYRDVLKEMKAKASTVKALLSVEERACKLTPPHSAKSKFGYGAADVRSLSRAVNHIR 124
Db 61 DDHYRDVLKEMKAKASTVKALLSVEERACKLTPPHSAKSKFGYGAADVRSLSRAVNHIIH 120
QY 125 SVWKDLLEDDTPIQTTIMAKNEVFCVQPEKGRKPARLIIVPDDLGVRCRMALYDVVS 184
Db 121 SVWKDLLEDDTPIQTTIMAKNEVFCVQPEKGRKPARLIIVPDDLGVRCRMALYDVVS 180
QY 185 TLPQAVMGSSYGFQYSPQORVEFLVNTWKAKCPMGFSYDTRCPDSTVTENDIRVEESIY 244
Db 181 TLPQAVMGSSYGFQYSPQORVEFLVNTWKAKCPMGFSYDTRCPDSTVTENDIRVEESIY 240
QY 245 OCCDLAPARQAIRSLTERLYVGGPMTNSKQNCYRRCRASGLVLTSCGNTLTCLYKAA 304
Db 241 OCCDLAPARQAIRSLTERLYVGGPMTNSKQNCYRRCRASGLVLTSCGNTLTCLYKAS 300
QY 305 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
Db 301 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 360
QY 365 ELITSCSSNVSAHDASGRVYILTRDTPVPLARAAMETARHTPVNSWLGNIIMYAPTILW 424
Db 361 ELITSCSSNVSAHDASGRVYILTRDTPVPLARAAMETARHTPVNSWLGNIIMYAPTILW 420
QY 425 ARMLIMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIEHLHGLSAFSLHSYSPG 484
Db 421 ARMLIMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIEHLHGLSAFSLHSYSPG 480
QY 485 EINRVASCLRLKGLVPPPLRVWRHRSVRKALLSQGGRAAICGKYLFWNAVTKLKLTPIP 544
Db 481 EINRVASCLRLKGLVPPPLRVWRHRSVRKALLSQGGRAATCGKYLFWNAVTKLKLTPIP 540
QY 545 AASRLDLSGWFVAGYSGGDIYHSLSRAPRPFMLCLLLLSVGVGIYLLPNR 595
Db 541 AASRLDLSGWFVAGYSGGDIYHSLSRAPRPFMLCLLLLSVGVGIYLLPNR 591
```

RESULT 4

```
US-10-085-476-2
; Sequence 2, Application US/10085476
; Publication No. US20020164722A1
; GENERAL INFORMATION:
; APPLICANT: De Francesco, Raffaele
; APPLICANT: Tomei, Lucia
; APPLICANT: Behrens, Sven-Erik
; TITLE OF INVENTION: METHOD FOR REPRODUCING IN VITRO THE
; TITLE OF INVENTION: RNA-DEPENDENT RNA POLYMERASE AND TERMINAL NUCLEOTIDYL
; TITLE OF INVENTION: TRANSFERASE ACTIVITIES ENCODED BY HEPATITIS C VIRUS (HCV)
; FILE REFERENCE: IT0002PCA
; CURRENT APPLICATION NUMBER: US/10/085,476
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: 08/952,981
; PRIOR FILING DATE: 1998-03-23
; PRIOR APPLICATION NUMBER: PCT/IT96/00106
; PRIOR FILING DATE: 1996-05-24
; PRIOR APPLICATION NUMBER: RM95A000343
; PRIOR FILING DATE: 1995-05-25
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2201
; TYPE: PRT
; ORGANISM: cDNA clone pCD (38-9.4)
US-10-085-476-2
```

```
Query Match 95.2%; Score 3016; DB 13; Length 2201;
Best Local Similarity 96.6%; Pred. No. 1.7e-265;
Matches 571; Conservative 11; Mismatches 9; Indels 0; Gaps 0;
```

```
QY 5 SMSYTWTCALITPCAABESKLPINALSNSLLRHNLVYTTSSRSASLRQKKVTFDRLOVL 64
Db 1611 SMSYTWTCALITPCAABESKLPINALSNSLLRHNLVYTTSSRSASLRQKKVTFDRLOVL 1670
QY 65 DDHYRDVLKEMKAKASTVKALLSVEERACKLTPPHSAKSKFGYGAADVRSLSRAVNHIR 124
Db 1671 DDHYRDVLKEMKAKASTVKALLSVEERACKLTPPHSAKSKFGYGAADVRSLSRAVNHIIH 1730
QY 125 SVWKDLLEDDTPIQTTIMAKNEVFCVQPEKGRKPARLIIVPDDLGVRCRMALYDVVS 184
Db 1731 SVWKDLLEDDTPIQTTIMAKNEVFCVQPEKGRKPARLIIVPDDLGVRCRMALYDVVS 1790
QY 185 TLPQAVMGSSYGFQYSPQORVEFLVNTWKAKCPMGFSYDTRCPDSTVTENDIRVEESIY 244
Db 1791 TLPQAVMGSSYGFQYSPQORVEFLVNTWKAKCPMGFSYDTRCPDSTVTENDIRVEESIY 1850
QY 245 OCCDLAPARQAIRSLTERLYVGGPMTNSKQNCYRRCRASGLVLTSCGNTLTCLYKAA 304
Db 1951 OCCDLAPARQAIRSLTERLYVGGPMTNSKQNCYRRCRASGLVLTSCGNTLTCLYKAS 1910
QY 305 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
Db 1911 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 1970
QY 365 ELITSCSSNVSAHDASGRVYILTRDTPVPLARAAMETARHTPVNSWLGNIIMYAPTILW 424
Db 1971 ELITSCSSNVSAHDASGRVYILTRDTPVPLARAAMETARHTPVNSWLGNIIMYAPTILW 2030
QY 425 ARMLIMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIEHLHGLSAFSLHSYSPG 484
Db 2031 ARMLIMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIEHLHGLSAFSLHSYSPG 2090
QY 485 EINRVASCLRLKGLVPPPLRVWRHRSVRKALLSQGGRAAICGKYLFWNAVTKLKLTPIP 544
Db 2091 EINRVASCLRLKGLVPPPLRVWRHRSVRKALLSQGGRAATCGKYLFWNAVTKLKLTPIP 2150
QY 545 AASRLDLSGWFVAGYSGGDIYHSLSRAPRPFMLCLLLLSVGVGIYLLPNR 595
Db 2151 AASRLDLSGWFVAGYSGGDIYHSLSRAPRPFMLCLLLLSVGVGIYLLPNR 2201
```

RESULT 5

```
US-10-241-872-23
; Sequence 23, Application US/10241872
; Publication No. US20030152915A1
; GENERAL INFORMATION:
; APPLICANT: Hagedorn, Curt H.
; TITLE OF INVENTION: Recombinant Hepatitis C Virus RNA Replicase
; FILE REFERENCE: 76-95C
; CURRENT APPLICATION NUMBER: US/10/241,872
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/597,877
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 60/004,383
; PRIOR FILING DATE: 1995-09-27
; PRIOR APPLICATION NUMBER: 08/722,806
; PRIOR FILING DATE: 1996-09-27
; PRIOR APPLICATION NUMBER: 09/337,028
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 23
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-241-872-23
```

```
Query Match 95.2%; Score 3014; DB 14; Length 593;
Best Local Similarity 96.3%; Pred. No. 3.8e-266;
Matches 569; Conservative 12; Mismatches 10; Indels 0; Gaps 0;
QY 5 SMSYTWTCALITPCAABESKLPINALSNSLLRHNLVYTTSSRSASLRQKKVTFDRLOVL 64
```

Db 3 SMSYTWGALITPCAAEESKLPINALNSLRRHNHNVYATTSRSASLRQKKVTFDRQLVL 62
QY 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYKAKDVRSLSRAVNHIR 124
Db 63 DDHYQDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYKAKDVRSLSRAVNHIR 122
QY 125 SVWKDLLEDTDTPQTITIMAKNEVFCVQPEKGGKRPARLIVFPDGLGVRVCEKMAKYDVVS 184
Db 123 SVWKDLLEDTDTPQTITIMAKNEVFCVQPEKGGKRPARLIVFPDGLGVRVCEKMAKYDVVS 182
QY 185 TLPOAVMGSSYGFQYSPQKRVFELVNTWKAKCPMGFSYDTRCFDSTVTENDIRVESIY 244
Db 183 TLPOAVMGSSYGFQYSPQKRVFELVNTWKAKCPMGFSYDTRCFDSTVTENDIRVESIY 242
QY 245 QCCDLAPARQAISLTERLYVGGPMTNSKQNGCYRRCRASGVLTSSCGNTLTCYLKAA 304
Db 243 QCCDLAPARQAISLTERLYVGGPMTNSKQNGCYRRCRASGVLTSSCGNTLTCYLKAA 302
QY 305 AACRAAKLQDCTMLVNGDDLVIWICESAGTQDAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
Db 303 AACRAAKLQDCTMLVNGDDLVIWICESAGTQDAASLRVFTTEAMTRYSAAPPDPPQPEYDL 362
QY 365 ELITSCSSNVSAHDASGRVYVYLTDRPTVPLARAAMETARHTPVNSWLGNIIMYAPTLW 424
Db 363 ELITSCSSNVSAHDASGRVYVYLTDRPTVPLARAAMETARHTPVNSWLGNIIMYAPTLW 422
QY 425 ARMLTMTHFFSILLAOLEKALDCQIYGACYSIEPLDLPQIIRLHGLSFAFSLSHSPG 484
Db 423 ARMLTMTHFFSILLAOLEKALDCQIYGACYSIEPLDLPQIIRLHGLSFAFSLSHSPG 482
QY 485 EIRNVASCLRLKGLVPPPLRVWHRARSVRAKLLSQGGRAAICGKYLEFNWAVTKLKLTPIP 544
Db 483 EIRNVASCLRLKGLVPPPLRVWHRARSVRAKLLSQGGRAAICGKYLEFNWAVTKLKLTPIP 542
QY 545 AASRLDLSGMFVAGYSGGDIYHLSLRARPRFWMCLLLLSVGVGIYLLPNR 595
Db 543 AASQLDLSGMFVAGYSGGDIYHLSLRARPRFWMCLLLLSVGVGIYLLPNR 593

RESULT 6

US-10-492-178-6
; Sequence 6, Application US/10492178
; Publication No. US20040247615A1
; GENERAL INFORMATION:
; APPLICANT: Emimi, Emilio A.
; APPLICANT: Kaslow, David C.
; APPLICANT: Bett, Andrew J.
; APPLICANT: Shiver, John W.
; APPLICANT: Nicosia, Alfredo
; APPLICANT: Lahm, Armin
; APPLICANT: Luzzago, Alessandra
; APPLICANT: Cortese, Riccardo
; APPLICANT: Colloca, Stefano
; TITLE OF INVENTION: HEPATITIS C VIRUS VACCINE
; FILE REFERENCE: ITR0015YP
; CURRENT APPLICATION NUMBER: US/10/492,178
; PRIOR FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: PCT/US02/32512
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: 60/363,774
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/328,655
; PRIOR FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1985
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: NS sequence

US-10-492-178-6
QY 5 SMSYTWGALITPCAAEESKLPINALNSLRRHNHNVYATTSRSASLRQKKVTFDRQLVL 64
Db 1 SMSYTWGALITPCAAEESKLPINALNSLRRHNHNVYATTSRSASLRQKKVTFDRQLVL 60

Query Match 95.1%; Score 3012; DB 16; Length 1985;
Best Local Similarity 96.4%; Pred. No. 3.4e-265;
Matches 570; Conservative 12; Mismatches 9; Indels 0; Gaps 0;
QY 5 SMSYTWGALITPCAAEESKLPINALNSLRRHNHNVYATTSRSASLRQKKVTFDRQLVL 64
Db 1395 SMSYTWGALITPCAAEESKLPINALNSLRRHNHNVYATTSRSASLRQKKVTFDRQLVL 1454
QY 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYKAKDVRSLSRAVNHIR 124
Db 1455 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYKAKDVRSLSRAVNHIR 1514
QY 125 SVWKDLLEDTDTPQTITIMAKNEVFCVQPEKGGKRPARLIVFPDGLGVRVCEKMAKYDVVS 184
Db 1515 SVWKDLLEDTDTPQTITIMAKNEVFCVQPEKGGKRPARLIVFPDGLGVRVCEKMAKYDVVS 1574
QY 185 TLPOAVMGSSYGFQYSPQKRVFELVNTWKAKCPMGFSYDTRCFDSTVTENDIRVESIY 244
Db 1575 TLPOAVMGSSYGFQYSPQKRVFELVNTWKAKCPMGFSYDTRCFDSTVTENDIRVESIY 1634
QY 245 QCCDLAPARQAISLTERLYVGGPMTNSKQNGCYRRCRASGVLTSSCGNTLTCYLKAA 304
Db 1635 QCCDLAPARQAISLTERLYVGGPMTNSKQNGCYRRCRASGVLTSSCGNTLTCYLKAA 1694
QY 305 AACRAAKLQDCTMLVNGDDLVIWICESAGTQDAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
Db 1695 AACRAAKLQDCTMLVNGDDLVIWICESAGTQDAASLRVFTTEAMTRYSAAPPDPPQPEYDL 1754
QY 365 ELITSCSSNVSAHDASGRVYVYLTDRPTVPLARAAMETARHTPVNSWLGNIIMYAPTLW 424
Db 1755 ELITSCSSNVSAHDASGRVYVYLTDRPTVPLARAAMETARHTPVNSWLGNIIMYAPTLW 1814
QY 425 ARMLTMTHFFSILLAOLEKALDCQIYGACYSIEPLDLPQIIRLHGLSFAFSLSHSPG 484
Db 1815 ARMLTMTHFFSILLAOLEKALDCQIYGACYSIEPLDLPQIIRLHGLSFAFSLSHSPG 1874
QY 485 EIRNVASCLRLKGLVPPPLRVWHRARSVRAKLLSQGGRAAICGKYLEFNWAVTKLKLTPIP 544
Db 1875 EIRNVASCLRLKGLVPPPLRVWHRARSVRAKLLSQGGRAAICGKYLEFNWAVTKLKLTPIP 1934
QY 545 AASRLDLSGMFVAGYSGGDIYHLSLRARPRFWMCLLLLSVGVGIYLLPNR 595
Db 1935 AASQLDLSGMFVAGYSGGDIYHLSLRARPRFWMCLLLLSVGVGIYLLPNR 1985

RESULT 7

US-10-471-164-5
; Sequence 5, Application US/10471164
; Publication No. US20040110126A1
; GENERAL INFORMATION:
; APPLICANT: George KUKOLJ and Ginette MCKERCHER
; TITLE OF INVENTION: HCV POLYMERASE INHIBITOR ASSAY
; FILE REFERENCE: 13/094
; CURRENT APPLICATION NUMBER: US/10/471,164
; CURRENT FILING DATE: 2003-09-05
; PRIOR APPLICATION NUMBER: 60/274,374
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 591
; TYPE: PRT
; ORGANISM: HCV
; US-10-471-164-5

Query Match 95.1%; Score 3011; DB 16; Length 591;
Best Local Similarity 95.9%; Pred. No. 7.1e-266;
Matches 567; Conservative 18; Mismatches 6; Indels 0; Gaps 0;

QY 5 SMSYTWGALITPCAAEESKLPINALNSLRRHNHNVYATTSRSASLRQKKVTFDRQLVL 64
Db 1 SMSYTWGALITPCAAEESKLPINALNSLRRHNHNVYATTSRSASLRQKKVTFDRQLVL 60

QY 65 DDHYRDVLEKMAKASTVAKALLSVEEACKLTPPHSAKSKFGYGAADVRSLSRAVNHIR 124
Db 61 DDHYRDVLEKMAKASTVAKALLSVEEACKLTPPHSAKSKFGYGAADVRSLSRAVNHIR 120
QY 125 SVWKOLLEDTPTIOTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 184
Db 121 SVWKOLLEDTPTIOTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 180
QY 185 TLPOAVMGSSYGFQYSPKORVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
Db 181 TLPOAVMGSSYGFQYSPKORVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 240
QY 245 OCCDLAPARQAIRSLTERLYVGGPMNTSKGONCGYRRCRASGVLTTCGNTLTCTYLKAA 304
Db 241 OCCDLAPARQAIRSLTERLYVGGPMNTSKGONCGYRRCRASGVLTTCGNTLTCTYLKAS 300
QY 305 AACRAAKLQDCTMLVNGDDLVI CESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
Db 301 AACRAAKLQDCTMLVNGDDLVI CESAGTQEDAAANLRVFTTEAMTRYSAAPPDPPQPEYDL 360
QY 365 ELITSCSSNVSAHDASGRVYVLTTRDPTVPLARAAMETARHTPNSWMLGNIIMYAPTILW 424
Db 361 ELITSCSSNVSAHDASGRVYVLTTRDPTVPLARAAMETARHTPNSWMLGNIIMYAPTILW 420
QY 425 ARMILMTHFFSILLAQOELEKALDCQIYGACYSIEPLDLPQIIRLHGLSAPLSHSYSPG 484
Db 421 ARMVLMTHFFSILLAQOELEKALDCQIYGACYSIEPLDLPQIIRLHGLSAPLSHSYSPG 480
QY 485 EINRVASCLRLKGLVPPPLRVWRHRARSVRKLLSQGGRAATCGKYLFFNNAVTRKLTPTIP 544
Db 481 EINRVASCLRLKGLVPPPLRVWRHRARSVRKLLSQGGRAATCGKYLFFNNAVTRKLTPTIP 540
QY 545 AASRLDLSGFWVAGYSGGDIYHLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
Db 541 AASRLDLSGFWVAGYSGGDIYHLSRARPRWFMLCLLLLSVGVGIYLLPNR 591

RESULT 8

US-09-898-297-1
; Sequence 1, Application US/09898297
; Publication No. US20020065418A1
; GENERAL INFORMATION:
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: VIRAL POLYMERASE INHIBITORS
; CURRENT APPLICATION NUMBER: US/09/898,297
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/216,084
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 60/274,374
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/281,434
; PRIOR FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 621
; TYPE: PRT
; ORGANISM: Hepatitis C Virus

Query Match 95.1%; Score 3011; DB 9; Length 621;
Best Local Similarity 95.9%; Pred. No. 7,6e-266;
Matches 567; Conservative 18; Mismatches 6; Indels 0; Gaps 0;

QY 5 SMSYTWTCALITPCAAEESKLPINALSNSLRHNLVYSTTSRSASLRQKKVTFDRLOVL 64
Db 31 SMSYTWTCALITPCAAEESKLPINALSNSLRHNLVYSTTSRSASLRQKKVTFDRLOVL 90
QY 65 DDHYRDVLEKMAKASTVAKALLSVEEACKLTPPHSAKSKFGYGAADVRSLSRAVNHIR 124
Db 91 DDHYRDVLEKMAKASTVAKALLSVEEACKLTPPHSAKSKFGYGAADVRSLSRAVNHIR 150

QY 125 SVWKOLLEDTPTIOTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 184
Db 151 SVWKOLLEDTPTIOTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 210
QY 185 TLPOAVMGSSYGFQYSPKORVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
Db 211 TLPOAVMGSSYGFQYSPKORVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 270
QY 245 OCCDLAPARQAIRSLTERLYVGGPMNTSKGONCGYRRCRASGVLTTCGNTLTCTYLKAA 304
Db 271 OCCDLAPARQAIRSLTERLYVGGPMNTSKGONCGYRRCRASGVLTTCGNTLTCTYLKAS 330
QY 305 AACRAAKLQDCTMLVNGDDLVI CESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
Db 331 AACRAAKLQDCTMLVNGDDLVI CESAGTQEDAAANLRVFTTEAMTRYSAAPPDPPQPEYDL 390
QY 365 ELITSCSSNVSAHDASGRVYVLTTRDPTVPLARAAMETARHTPNSWMLGNIIMYAPTILW 424
Db 391 ELITSCSSNVSAHDASGRVYVLTTRDPTVPLARAAMETARHTPNSWMLGNIIMYAPTILW 450
QY 425 ARMILMTHFFSILLAQOELEKALDCQIYGACYSIEPLDLPQIIRLHGLSAPLSHSYSPG 484
Db 451 ARMVLMTHFFSILLAQOELEKALDCQIYGACYSIEPLDLPQIIRLHGLSAPLSHSYSPG 510
QY 485 EINRVASCLRLKGLVPPPLRVWRHRARSVRKLLSQGGRAATCGKYLFFNNAVTRKLTPTIP 544
Db 511 EINRVASCLRLKGLVPPPLRVWRHRARSVRKLLSQGGRAATCGKYLFFNNAVTRKLTPTIP 570
QY 545 AASRLDLSGFWVAGYSGGDIYHLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
Db 571 AASRLDLSGFWVAGYSGGDIYHLSRARPRWFMLCLLLLSVGVGIYLLPNR 621

RESULT 9

US-10-198-680A-1
; Sequence 1, Application US/10198680A
; Publication No. US20030176433A1
; GENERAL INFORMATION:
; APPLICANT: Boehringer Ingelheim (Canada) Ltd.
; TITLE OF INVENTION: Viral Polymerase Inhibitors
; FILE REFERENCE: 13/095
; CURRENT APPLICATION NUMBER: US/10/198,680A
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/338,061
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 60/307,674
; PRIOR FILING DATE: 2001-07-25
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 621
; TYPE: PRT
; ORGANISM: HCV NS5B

Query Match 95.1%; Score 3011; DB 14; Length 621;
Best Local Similarity 95.9%; Pred. No. 7,6e-266;
Matches 567; Conservative 18; Mismatches 6; Indels 0; Gaps 0;

QY 5 SMSYTWTCALITPCAAEESKLPINALSNSLRHNLVYSTTSRSASLRQKKVTFDRLOVL 64
Db 31 SMSYTWTCALITPCAAEESKLPINALSNSLRHNLVYSTTSRSASLRQKKVTFDRLOVL 90
QY 65 DDHYRDVLEKMAKASTVAKALLSVEEACKLTPPHSAKSKFGYGAADVRSLSRAVNHIR 124
Db 91 DDHYRDVLEKMAKASTVAKALLSVEEACKLTPPHSAKSKFGYGAADVRSLSRAVNHIR 150
QY 125 SVWKOLLEDTPTIOTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 184
Db 151 SVWKOLLEDTPTIOTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 210
QY 185 TLPOAVMGSSYGFQYSPKORVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244

Db 211 TLPOAVMGSSYGFQSPKQVFEFLVNAWKKCPMGFSYDTRCFDSTVTESDIRVEESIY 270
Qy 245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKQNCYRRCRASGLVTTSCGNTLTCLYKAA 304
Db 271 QCCDLAPEARQAIRSLTERLYVGGPMTNSKQNCYRRCRASGLVTTSCGNTLTCLYKAS 330
Qy 305 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAALRVFTTEAMTRYSAAPPDLPQPEYDL 364
Db 331 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAALRVFTTEAMTRYSAAPPDLPQPEYDL 390
Qy 365 ELITSCSSNVSAHDASGRVYVLTDRPTVPLARAAMETARHTPVNSWLGNIIWVAPTLW 424
Db 391 ELITSCSSNVSAHDASGRVYVLTDRPTVPLARAAMETARHTPVNSWLGNIIWVAPTLW 450
Qy 425 ARMILMTHFFSILLAOBLEKALDCQIYGACYSIEPLDLPOIIEHLHGLSAFSLHSYSPG 484
Db 451 ARMVLMTHFFSILLAOBLEKALDCQIYGACYSIEPLDLPOIIEHLHGLSAFSLHSYSPG 510
Qy 485 EINRVASCLRLKLGVPPLRVWHRARSVRAKLLSQGGRAAICGKYLFWNAVRTKLTPIIP 544
Db 511 EINRVASCLRLKLGVPPLRVWHRARSVRAKLLSQGGRAAICGKYLFWNAVRTKLTPIIP 570
Qy 545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMCLLLLLSVGVGIYLLPNR 595
Db 571 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMCLLLLLSVGVGIYLLPNR 621

RESULT 10

US-10-238-282-1
; Sequence 1, Application US/10238282
; Publication No. US20030232816A1
; GENERAL INFORMATION:
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: VIRAL POLYMERASE INHIBITORS
; FILE REFERENCE: 13/079
; CURRENT APPLICATION NUMBER: US/10/238,282
; PRIOR FILING DATE: 2002-09-10
; PRIOR APPLICATION NUMBER: 60/216,084
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 60/274,374
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/281,434
; PRIOR FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 621
; TYPE: PRT
; ORGANISM: Hepatitis C Virus
US-10-238-282-1

Query Match 95.1%; Score 3011; DB 15; Length 621;
Best Local Similarity 95.9%; Pred. No. 7.6e-266;
Matches 567; Conservative 18; Mismatches 6; Indels 0; Gaps 0;
Qy 5 SMSYTWTCALITPCAABESKLPINALNSLRHNLVYSTTSRSASLRQKVTDFRLOVL 64
Db 31 SMSYTWTCALITPCAABESKLPINALNSLRHNLVYSTTSRSASLRQKVTDFRLOVL 90
Qy 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAQVRSLSRAVNHIR 124
Db 91 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAQVRSLSRAVNHIR 150
Qy 125 SVWKDLLEDTPPTQTTIMAKNEVFCVQPEKGRKPARLI VFPDLGVRVCCKMALYDVWS 184
Db 151 SVWKDLLEDTPPTQTTIMAKNEVFCVQPEKGRKPARLI VFPDLGVRVCCKMALYDVWS 210
Qy 185 TLPOAVMGSSYGFQSPKQVFEFLVNAWKKCPMGFSYDTRCFDSTVTESDIRVEESIY 244
Db 211 TLPOAVMGSSYGFQSPKQVFEFLVNAWKKCPMGFSYDTRCFDSTVTESDIRVEESIY 270
Qy 245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKQNCYRRCRASGLVTTSCGNTLTCLYKAA 304
Db 271 QCCDLAPEARQAIRSLTERLYVGGPMTNSKQNCYRRCRASGLVTTSCGNTLTCLYKAS 330
Qy 305 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAALRVFTTEAMTRYSAAPPDLPQPEYDL 364
Db 331 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAALRVFTTEAMTRYSAAPPDLPQPEYDL 390

Db 271 QCCDLAPEARQAIRSLTERLYVGGPMTNSKQNCYRRCRASGLVTTSCGNTLTCLYKAS 330
Qy 305 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAALRVFTTEAMTRYSAAPPDLPQPEYDL 364
Db 331 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAALRVFTTEAMTRYSAAPPDLPQPEYDL 390
Qy 365 ELITSCSSNVSAHDASGRVYVLTDRPTVPLARAAMETARHTPVNSWLGNIIWVAPTLW 424
Db 391 ELITSCSSNVSAHDASGRVYVLTDRPTVPLARAAMETARHTPVNSWLGNIIWVAPTLW 450
Qy 425 ARMILMTHFFSILLAOBLEKALDCQIYGACYSIEPLDLPOIIEHLHGLSAFSLHSYSPG 484
Db 451 ARMVLMTHFFSILLAOBLEKALDCQIYGACYSIEPLDLPOIIEHLHGLSAFSLHSYSPG 510
Qy 485 EINRVASCLRLKLGVPPLRVWHRARSVRAKLLSQGGRAAICGKYLFWNAVRTKLTPIIP 544
Db 511 EINRVASCLRLKLGVPPLRVWHRARSVRAKLLSQGGRAAICGKYLFWNAVRTKLTPIIP 570
Qy 545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMCLLLLLSVGVGIYLLPNR 595
Db 571 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMCLLLLLSVGVGIYLLPNR 621

RESULT 11

US-10-198-259A-1
; Sequence 1, Application US/10198259A
; Publication No. US20030236251A1
; GENERAL INFORMATION:
; APPLICANT: Boehringer Ingelheim (Canada) Ltd.
; TITLE OF INVENTION: Viral Polymerase Inhibitors
; FILE REFERENCE: 13/089
; CURRENT APPLICATION NUMBER: US/10/198,259A
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,669
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/338,324
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 621
; TYPE: PRT
; ORGANISM: HCV NS5B
US-10-198-259A-1

Query Match 95.1%; Score 3011; DB 15; Length 621;
Best Local Similarity 95.9%; Pred. No. 7.6e-266;
Matches 567; Conservative 18; Mismatches 6; Indels 0; Gaps 0;

Qy 5 SMSYTWTCALITPCAABESKLPINALNSLRHNLVYSTTSRSASLRQKVTDFRLOVL 64
Db 31 SMSYTWTCALITPCAABESKLPINALNSLRHNLVYSTTSRSASLRQKVTDFRLOVL 90
Qy 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAQVRSLSRAVNHIR 124
Db 91 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAQVRSLSRAVNHIR 150
Qy 125 SVWKDLLEDTPPTQTTIMAKNEVFCVQPEKGRKPARLI VFPDLGVRVCCKMALYDVWS 184
Db 151 SVWKDLLEDTPPTQTTIMAKNEVFCVQPEKGRKPARLI VFPDLGVRVCCKMALYDVWS 210
Qy 185 TLPOAVMGSSYGFQSPKQVFEFLVNAWKKCPMGFSYDTRCFDSTVTESDIRVEESIY 244
Db 211 TLPOAVMGSSYGFQSPKQVFEFLVNAWKKCPMGFSYDTRCFDSTVTESDIRVEESIY 270
Qy 245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKQNCYRRCRASGLVTTSCGNTLTCLYKAA 304
Db 271 QCCDLAPEARQAIRSLTERLYVGGPMTNSKQNCYRRCRASGLVTTSCGNTLTCLYKAS 330
Qy 305 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAALRVFTTEAMTRYSAAPPDLPQPEYDL 364
Db 331 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAALRVFTTEAMTRYSAAPPDLPQPEYDL 390

Qy	365	ELITSCSSNSVAVHDASGRVYLLTRDTPVLABAAWETARHTPNSWLGNIIMVAPTILW	424
Db	391	ELITSCSSNSVAVHDASGRVYLLTRDTPVLABAAWETARHTPNSWLGNIIMVAPTILW	450
Qy	425	ARMILMTHFPFISILLAQEQLEKALDCQIYGACVSYSTIEPLDLPTQIERLHGLSAPLSHSYSPG	484
Db	451	ARMVLMTHFPFISILLAQEQLEKALDCQIYGACVSYSTIEPLDLPTQIERLHGLSAPLSHSYSPG	510
Qy	485	EINRVASCLRLKGVPPRLVVRHRRARSVRKLLSQGGRAAICCKYLFNNAVTKLKLTIP	544
Db	511	EINRVASCLRLKGVPPRLVVRHRRARSVRKLLSQGGRAATCGKYLFNNAVTKLKLTIP	570
Qy	545	AASRLDLSGWFVAGYSGGDIYHSLSRARPWFMLCLLLLSVGVGIGYLLPNR	595
Db	571	AASRLDLSGWFVAGYSGGDIYHSLSRARPWFMLCLLLLSVGVGIGYLLPNR	621

```

RESULT 12
US-10-198-384A-1
; Sequence 1, Application US/10198384A
; Publication No. US20040024190A1
; GENERAL INFORMATION:
; APPLICANT: Boehringer Ingelheim (Canada) Ltd.
; TITLE OF INVENTION: Viral Polymerase inhibitors
; FILE REFERENCE: 13/090
; CURRENT APPLICATION NUMBER: US/10/198,384A
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/338,061
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 60/307,674
; PRIOR FILING DATE: 2001-07-25
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 621
; TYPE: PrT
; ORGANISM: HCV NS5B
US-10-198-384A-1

```

Query Match	95.1%	Score 3011;	DB 15;	Length 621;
Best Local Similarity	95.9%;	Pred. No. 7.6e-266;		
Matches 567;	Conservative 18;	Mismatches 6;	Indels 0;	Gaps 0;
QY	5	SMSYTTTGALITTC	AAEESKLPINALSNSILRRHNLVYSTTSRSASLURQKQVTFDRQLVL	64
Db	31	SMSYTTTGALITTC	AAEESQIPINALSNSILVRHNMVYSTTSRAALRQKQVTFDRQLVL	90
QY	65	DDHYRDVLKEMKAKASTV	KAKLSVBEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR	124
Db	91	DDHYRDVLKEMKAKASTV	KAKLSVBEACKLTPPHSAKSKFGYGAKDVRNLSSKAVDHIR	150
QY	125	SVMKDLLEDTDTP	IQTTIMAKNEVFCVQPEKGGKPARLIVFPDGLVYRVECKMALYDWS	184
Db	151	SVMKDLLEDTET	ITDITIMAKNEVFCVQPEKGGKPARLIVFPDGLVYRVECKMALYDWS	210
QY	185	TLPOAVWGSSYGFQYS	PQQRVEFLNWTWAKKCPMGFSYDTRCFDSTVTETENDLRVESIY	244
Db	211	TLPOAVWGSSYGFQYS	PQQRVEFLVNAWKSKCPMGFSYDTRCFDSTVTETENDLRVESIY	270
QY	245	QCDDLAPEARQAI	RSLTERLYVGGPMTNSXGQNGCYRRCRASGVLTTSCGNLTTCYLKAA	304
Db	271	QCDDLAPEARQAI	KSALTERLYIGGPLTNSXGQNGCYRRCRASGVLTTSCGNLTTCYLKAS	330
QY	305	AACRAAKLODCTML	VNGDDLWVTCESAGTQEDAAASLAVFTTEAMTRYISAPGDDPPQPYDL	364
Db	331	AACRAAKLODCTML	VNGDDLWVTCESAGTQEDANLAVFTTEAMTRYISAPGDLFQPEYDL	390
QY	365	ELITSCSSNVSV	AHDASGKRVYVLTDRDPTVPLARAAMETARHTPVNSWLGNIMYAPTLW	424
Db	391	ELITSCSSNVSV	AHDASGKRVYVLTDRDPTVPLARAAMETARHTPVNSWLGNIMYAPTLW	450
QY	425	ARMILMTHPFS	ILLAQSEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSPG	484

451	ARWVLTHTFSSILLAQBLEKALDCQYGYACYSYTEPLDLPQIIERLHGLSAPLSHSYSPG	510
Db		
485	EINRVASCLSKLGVPPPLRVWRHRRARSVRAKLLSQGGRAAI CKYILFNNAVRTKLKLTPIP	544
Qy		
511	EINRVASCLSKLGVPPPLRVWRHRRARSVRAKLLSQGGRAATCKYILFNNAVRTKLKLTPIP	570
Db		
545	AASRLDLSGWFVAGYSGGDIYHLSLSRAPRWFMLCLLLSVGVGVIYLLPNR	595
Qy		
571	AASRLDLSGWFVAGYSGGDIYHLSLSRAPRWFMLCLLLSVGVGVIYLLPNR	621
Db		

```

RESULT 13
US-10-471-164-1
; Sequence 1, Application US/10471164
; Publication No. US20040110126A1
; GENERAL INFORMATION:
; APPLICANT: George KUKOLJ and Ginette MCKERCHER
; TITLE OF INVENTION: HCV POLYMERASE INHIBITOR ASSAY.
; FILE REFERENCE: 13/094
; CURRENT APPLICATION NUMBER: US/10/471,164
; CURRENT FILING DATE: 2003-09-05
; PRIOR APPLICATION NUMBER: 60/274,374
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 621
; TYPE: PRT
; ORGANISM: HCV
US-10-471-164-1

```

		Query Match	95.1%; Score 3011; DB 16; Length 621;
		Best Local Similarity	95.9%; Pred. No. 7.6e-266;
		Matches 567; Conservative	18; Mismatches 6; Indels 0; Gaps 0;
QY	5	SMSYTTTGALITPCAAEESKLPINALSNSLIRHNLVYSTTSRSASLRQKKVTDFRLQL	64
DB	31	SMSYTTTGALITPCAAEESQLPINALSNSLVHRNMVYSTTSRAALRQKKVTDFRLQL	90
QY	65	DDHYRDVLKMKAKASTVKAKLLSVEACKLTTPHSKSKFEGYGAQDVRSLSRAVNHIR	124
DB	91	DDHYRDVLKMKAKASTVKAKLLSVEACKLTTPHSKSKFEGYGAQDVRNLSKKAVDHIR	150
QY	125	SVMKDLLEDTDPTQTIMAKNEVFCVQPEKGGRKPARIIVFPDGLGVRCCKMALYDVVS	184
DB	151	SVMKDLLEDTEPTDTIMAKNEVFCVQPEKGGRKPARIIVFPDGLGVRCCKMALYDVVS	210
QY	185	TLPQAVMGSSYGFGSPQRQVEFLUNTWAKKCMPMGFSYDFTRCFDSTVTENDIRVEESI	244
DB	211	TLPQAVMGSSYGFGSPQRQVEFLVNAWKCKCMPGFSDYDFTRCFDSTVTESDIRVEESI	270
QY	245	OCCDLAPEARQAIRSLTERLYVGGMPTNSKGONCYRRCRASGVLTTSCGNWTLTCYLKAA	304
DB	271	OCCDLAPEARQAIRKSILTERLYIGGPLTNSKGONCYRRCRASGVLTTSCGNWTLTCYLKAS	330
QY	305	AACRAAKLODCTMLVNGDDLVIWCESAGTOEDAAASLRVFTTEAMTRYSAAPPDPPPOPYDL	364
DB	331	AACRAAKLODCTMLVNGDDLVIWCESAGTOEDAAANLRVFTTEAMTRYSAAPPDGLPOPYDL	390
QY	365	ELITSCSNVVAHDASGKRVYVLTDRPTVPLARAAMETABHTPVNSWLGINIMVAPTLM	424
DB	391	ELITSCSNVVAHDASGKRVYVLTDRPTVPLARAAMETAKHTPINSWGNIIMTAPTLM	450
QY	425	ARMILMTHFFSIILAQEQLEKALDCQIYGACYSTEPLDPQOIERLHGLSAFSLHSYSPG	484
DB	451	ARMVLMTHFFSIILAQEQLEKALDCQIYGACYSTEPLDPQOIERLHGLSAFSLHSYSPG	510
QY	485	EINRVASCRLKVPPPLVRVWRHARSVRAKLISOGGRAATCGKYLFNWAVENTKLKTPIP	544
DB	511	EINRVASCRLKVPPPLVRVWRHARSVRAKULLSOGGRAATCGKYLFNWAVENTKLKTPIP	570
QY	545	AASRLDSLGMFWFVAGYSGGDTYHLSLSRAPRMFMLCLLLSYGVGYILLPNR	595

Db 571 AASRLDLSGWFVAGYNGGDIYHLSRARPRWFMLCLLLLSVGVGIYLLPNR 621

RESULT 14

US-10-851-710-1
; Sequence 1, Application US/10851710
; Publication No. US20040224955A1
; GENERAL INFORMATION:
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: VIRAL POLYMERASE INHIBITORS
; FILE REFERENCE: 13/079
; CURRENT APPLICATION NUMBER: US/10/851,710
; CURRENT FILING DATE: 2004-05-21
; PRIOR APPLICATION NUMBER: US/10/238,282
; PRIOR FILING DATE: 2002-09-10
; PRIOR APPLICATION NUMBER: 60/216,084
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 60/274,374
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/281,434
; PRIOR FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 621
; TYPE: PRT
; ORGANISM: Hepatitis C Virus
US-10-851-710-1

Query Match 95.1%; Score 3011; DB 16; Length 621;
Best Local Similarity 95.9%; Pred. No. 7.6e-266;
Matches 567; Conservative 18; Mismatches 6; Indels 0; Gaps 0;

Qy 5 SMSYTWTCALITPCAAEESKLPINALSNSLRHNLVYVTSRSASLRQKKVTFDRLOVL 64
Db 31 SMSYTWTCALITPCAAEESKLPINALSNSLRHNLVYVTSRSASLRQKKVTFDRLOVL 90

Qy 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAOVRSLSRAVNHIR 124
Db 91 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAOVRSLSRAVNHIR 150

Qy 125 SWKDLLEDTTPIQTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 184
Db 151 SWKDLLEDTTPIQTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 210

Qy 185 TLPOAVMGSSYGFQYSPKQRFVFLVNTWAKKCPMGFSYDTRCFDSTVTSNDIRVEESIY 244
Db 211 TLPOAVMGSSYGFQYSPKQRFVFLVNTWAKKCPMGFSYDTRCFDSTVTSNDIRVEESIY 270

Qy 245 QCCDLAPPEARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTSCGNTLTCYLKAA 304
Db 271 QCCDLAPPEARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTSCGNTLTCYLKAA 330

Qy 305 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAASLRVFTTEAMTRYSAPPDPPQPEYDL 364
Db 331 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAANLRFVFTTEAMTRYSAPPDLPQPEYDL 390

Qy 365 ELITSCSSNVSAHDASGRVYVLTDRPTVPLARAANWETARHTPVNSWLGNIIMVAPTLW 424
Db 391 ELITSCSSNVSAHDASGRVYVLTDRPTVPLARAANWETARHTPVNSWLGNIIMVAPTLW 450

Qy 425 ARMILMTHFFSILLAOEQLEKALDCQIYGACYSIEPLDLPOIIRLHGLSAFSLHSYSPG 484
Db 451 ARMVLMTHFFSILLAOEQLEKALDCQIYGACYSIEPLDLPOIIRLHGLSAFSLHSYSPG 510

Qy 485 EINRVASCLRLKGLVPPPLRVWRHARSVRAKLLSQGGRAAICGKYLFWAVRTKLTPIIP 544
Db 511 EINRVASCLRLKGLVPPPLRVWRHARSVRAKLLSQGGRAAICGKYLFWAVRTKLTPIIP 570

Qy 545 AASRLDLSGWFVAGYNGGDIYHLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
Db 571 AASRLDLSGWFVAGYNGGDIYHLSRARPRWFMLCLLLLSVGVGIYLLPNR 621

RESULT 15

US-09-838-386-12
; Sequence 12, Application US/09838386
; Patent No. US20010055756A1
; GENERAL INFORMATION:
; APPLICANT: Pellerin, Charles
; APPLICANT: Kukulj, George
; TITLE OF INVENTION: Internal De No. US20010055756A10 Initiation Sites of the HCV NS5B
; FILE REFERENCE: 1011.2180001
; CURRENT APPLICATION NUMBER: US/09/838,386
; CURRENT FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: US 60/198,793
; PRIOR FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 12
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Artificial Sequence
; NAME/KEY: misc feature
; OTHER INFORMATION: HTa5B polymerase
US-09-838-386-12

Query Match 95.1%; Score 3011; DB 9; Length 627;
Best Local Similarity 95.9%; Pred. No. 7.7e-266;
Matches 567; Conservative 18; Mismatches 6; Indels 0; Gaps 0;

Qy 5 SMSYTWTCALITPCAAEESKLPINALSNSLRHNLVYVTSRSASLRQKKVTFDRLOVL 64
Db 37 SMSYTWTCALITPCAAEESKLPINALSNSLRHNLVYVTSRSASLRQKKVTFDRLOVL 96

Qy 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAOVRSLSRAVNHIR 124
Db 97 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAOVRSLSRAVNHIR 156

Qy 125 SWKDLLEDTTPIQTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 184
Db 157 SWKDLLEDTTPIQTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 216

Qy 185 TLPOAVMGSSYGFQYSPKQRFVFLVNTWAKKCPMGFSYDTRCFDSTVTSNDIRVEESIY 244
Db 217 TLPOAVMGSSYGFQYSPKQRFVFLVNTWAKKCPMGFSYDTRCFDSTVTSNDIRVEESIY 276

Qy 245 QCCDLAPPEARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTSCGNTLTCYLKAA 304
Db 277 QCCDLAPPEARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTSCGNTLTCYLKAA 336

Qy 305 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAASLRVFTTEAMTRYSAPPDPPQPEYDL 364
Db 337 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAANLRFVFTTEAMTRYSAPPDLPQPEYDL 396

Qy 365 ELITSCSSNVSAHDASGRVYVLTDRPTVPLARAANWETARHTPVNSWLGNIIMVAPTLW 424
Db 397 ELITSCSSNVSAHDASGRVYVLTDRPTVPLARAANWETARHTPVNSWLGNIIMVAPTLW 456

Qy 425 ARMILMTHFFSILLAOEQLEKALDCQIYGACYSIEPLDLPOIIRLHGLSAFSLHSYSPG 484
Db 457 ARMVLMTHFFSILLAOEQLEKALDCQIYGACYSIEPLDLPOIIRLHGLSAFSLHSYSPG 516

Qy 485 EINRVASCLRLKGLVPPPLRVWRHARSVRAKLLSQGGRAAICGKYLFWAVRTKLTPIIP 544
Db 517 EINRVASCLRLKGLVPPPLRVWRHARSVRAKLLSQGGRAAICGKYLFWAVRTKLTPIIP 576

Qy 545 AASRLDLSGWFVAGYNGGDIYHLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
Db 577 AASRLDLSGWFVAGYNGGDIYHLSRARPRWFMLCLLLLSVGVGIYLLPNR 627

Search completed: September 22, 2005, 14:49:31
Job time : 79 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2005, 14:44:29 ; Search time 78 Seconds
(without alignments)
2875.911 Million cell updates/sec

Title: US-10-712-479-4

Perfect score: 3055

Sequence: 1 MASMSMTWTGALITPCAA.....GGDIYHLSLRARPRHHHHH 580

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 1964263

Minimum DB seq length: 0

Maximum DB seq length: 585

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3055	100.0	580	8	ADO19023
2	2951	96.6	576	6	AAO26417
3	2948	96.5	576	6	AAO26419
4	2944	96.4	576	6	AAO26418
5	2942	96.3	579	4	AAO60131
6	2936	96.1	578	5	AAO79558
7	2936	96.1	578	6	ABP71662
8	2879	94.2	576	6	ABP72214
9	2879	94.2	576	6	ADH74494
10	2846	93.2	567	7	ADM69007
11	2734.5	89.5	545	5	AAO79559
12	2279	74.6	462	2	AAO58592
13	2279	74.6	462	8	ADL17783
14	2144	70.2	547	2	AAO67630
15	1812	59.3	389	2	AAO29906
16	1801	59.0	389	2	AAO29904
17	1798	58.9	389	2	AAO29905
18	1798	58.9	389	2	AAO29876
19	1755	57.4	365	2	AAO38284
20	1645	53.8	365	2	AAO38283
21	1634	53.5	365	2	AAO67592
22	1630	53.4	365	2	AAO67590
23	1430	46.8	352	2	AAO50081
24	1430	46.8	352	8	ADJ81686
25	1391	45.5	300	5	ABB77283

26	1391	45.5	365	2	AAO38287	Aar38287	NANB hepa
27	1384	45.3	365	2	AAO38286	Aar38286	NANB hepa
28	1375	45.0	365	2	AAO38285	Aar38285	NANB hepa
29	1363	44.6	285	2	AAO29879	Aar29879	HCV NS4-N
30	1232	40.3	260	5	ABB77285	Abb77285	HCV bait
31	1178	38.6	270	2	AAO25896	Aar25896	HK21. 9/2
32	1177	38.5	247	2	AAO25874	Aar25874	HCV poly
33	1177	38.5	247	2	AAO41752	Aaw41752	Hepatitis
34	1175	38.5	239	2	AAO25869	Aar25869	HCV poly
35	1175	38.5	239	2	AAO41747	Aaw41747	Hepatitis
36	1175	38.5	253	2	AAO25891	Aar25891	HK16. 9/2
37	1158.5	37.9	251	2	AAO25895	Aar25895	HK20. 9/2
38	1154	37.8	224	2	AAO25873	Aar25873	HCV poly
39	1154	37.8	224	2	AAO41751	Aaw41751	Hepatitis
40	1129	37.0	496	2	AAO21568	Aar21568	HCV CKS-N
41	1124	36.8	496	2	AAO33635	Aar33635	HCV CKS-N
42	1124	36.8	496	2	AAO33597	Aar33597	HCV CKS-N
43	1124	36.8	496	2	AAO33577	Aar33577	HCV CKS-N
44	1124	36.8	496	4	AAO51375	Abw51375	HCV recom
45	1124	36.8	496	7	ABW01861	Abw01861	HCV-CKS-N

ALIGNMENTS

RESULT 1

ADO19023

ID ADO19023 standard; protein; 580 AA.

AC ADO19023;

DT 12-AUG-2004 (first entry)

DE C-terminally truncated HCV NS5B polymerase, Cdelta21 NS5B.

XX RNA polymerase activity; continuous-read assay;

KW Hepatitis C virus infection; HCV infection; polymerase; NS5B;

KW bone mineral disease; osteoporosis; carcinoma; cardiovascular disease;

KW diabetes; ocular disorder; renal dysfunction; lymphoma;

KW lymphoproliferative disorder; metabolic disorder; arthritis;

KW sleep disorder; thyroid disorder; antiinflammatory; hepatotrophic;

KW virucide; osteopathic; cytostatic; antidiabetic; ophthalmological;

KW nephrotropic; antiarthritic; enzyme.

OS Hepatitis C virus.

PN WO200404228-A2.

XX 27-MAY-2004.

XX 13-NOV-2003; 2003WO-US036465.

PR 13-NOV-2002; 2002US-0425981P.

XX (PFIZ) PFIZER INC.

PI Yagi Y, Sheets MP, Wells PA, Shelly JA, Poorman RA, Epps DB;

PI Morgan AG;

XX WPI; 2004-420337/39.

DR N-PSDB; ADO19022.

XX Detecting RNA polymerase activity in a continuous-read manner, useful in

PT treating osteoporosis, carcinomas, cardiovascular diseases, ocular

PT disorders or arthritis, by contacting an RNA polymerase with an

PT oligonucleotide template.

PS Claim 7; SEQ ID NO 4; 46pp; English.

XX The present invention relates to a method for detecting RNA polymerase

CC activity in a continuous-read manner. The method comprises contacting an

CC RNA polymerase with an oligonucleotide template in a reaction mixture

CC comprising an assay buffer, under conditions in which the RNA polymerase

CC is active, adding a fluorescent dye capable of binding double-stranded
 CC nucleic acid molecules to the reaction mixture, and measuring the
 CC fluorescence of the reaction mixture. The RNA polymerase is the Hepatitis
 CC C virus (HCV) polymerase, NS5B. The method is useful for treating HCV
 CC infection, bone mineral diseases like osteoporosis, carcinomas,
 CC cardiovascular diseases, diabetes, ocular disorders, renal dysfunction,
 CC lymphomas, lymphoproliferative disorders, metabolic disorders, arthritis,
 CC sleep disorders and thyroid disorders. the present sequence represents C-
 CC terminally truncated HCV NS5B polymerase (designated Cdelta21 NS5B).
 XX
 SQ Sequence 580 AA;

Query Match 100.0%; Score 3055; DB 8; Length 580;
 Best Local Similarity 100.0%; Pred. No. 7e-285;
 Matches 580; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MASMSYTTGALITPCAAESKLPINALNSLRHNLVYVTSRSASLRQKVTDFR 60
 Db 1 MASMSYTTGALITPCAAESKLPINALNSLRHNLVYVTSRSASLRQKVTDFR 60

Qy 61 LQVLDHVRDLKEMKASTVKAKLSVEEACKLTPPHSAKSGFYGAKDVRSLSSRAV 120
 Db 61 LQVLDHVRDLKEMKASTVKAKLSVEEACKLTPPHSAKSGFYGAKDVRSLSSRAV 120

Qy 121 NHIRSVKDLLEDTPITQTTIMAKNEVFCVQPEKGGKRPALIVFPDLGVRVCEKMAV 180
 Db 121 NHIRSVKDLLEDTPITQTTIMAKNEVFCVQPEKGGKRPALIVFPDLGVRVCEKMAV 180

Qy 181 DVVSTLPQAVMGSSYGFQYSPKQKRVFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVE 240
 Db 181 DVVSTLPQAVMGSSYGFQYSPKQKRVFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVE 240

Qy 241 ESIYQCCDLAPEARQAIRSLTERLVVGGPMTNSKQNGYRCRASGLVLTSCGNLTTCY 300
 Db 241 ESIYQCCDLAPEARQAIRSLTERLVVGGPMTNSKQNGYRCRASGLVLTSCGNLTTCY 300

Qy 301 LKAAACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAASLRVTEAMTRYSAAPGDPPOP 360
 Db 301 LKAAACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAASLRVTEAMTRYSAAPGDPPOP 360

Qy 361 EYDLELITSCSNVSVAHDSAGKRVYVLTDRPTVPLAARAAETARHTPVNSWLGNIIMYA 420
 Db 361 EYDLELITSCSNVSVAHDSAGKRVYVLTDRPTVPLAARAAETARHTPVNSWLGNIIMYA 420

Qy 421 PTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHS 480
 Db 421 PTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHS 480

Qy 481 YSPGEINRVSCLRLKGVPPPLRVVHRARSVRAKLSQGGRAAICGKYLFWNAVTKLTP 540
 Db 481 YSPGEINRVSCLRLKGVPPPLRVVHRARSVRAKLSQGGRAAICGKYLFWNAVTKLTP 540

Qy 541 TPIPAASRLDLSGMFWAGVSGGDIYHSLSRARPRHHHHH 580
 Db 541 TPIPAASRLDLSGMFWAGVSGGDIYHSLSRARPRHHHHH 580

RESULT 2
 AAO26417
 ID AAO26417 standard; protein; 576 AA.
 XX
 AC AAO26417;
 XX
 DT 30-JAN-2003 (first entry)
 XX
 DE Mutant RNA polymerase MUT-2 protein.
 XX
 KW Antiviral; hepatitis C virus; HCV; NS5B RNA polymerase.
 XX
 OS Unidentified.
 XX
 PN EP125628-A2.
 XX

PD 13-NOV-2002.
 XX 07-MAY-2002; 2002EP-00009387.
 XX 10-MAY-2001; 2001US-0289829P.
 PR (AGOU-) AGOURON PHARM INC.
 PA Love RA, Yu X, Diehl W, Hickey MJ, Parge HE, Gao J, Fuhrman S;
 FI WPI; 2003-031968/03.
 DR
 XX
 XX Isolated nucleic acid encoding a mutant hepatitis C virus (HCV) NS5B RNA
 PT polymerase, useful for producing the mutant RNA polymerase which can be
 FT used for identifying inhibitors that can treat HCV infection.
 PT
 XX
 PS Claim 7; Page; 32pp; English.
 CC The invention relates to an isolated, purified nucleic acid molecule,
 CC which encodes a mutant hepatitis C virus (HCV) NS5B RNA polymerase, where
 CC a point mutation occurs at least at one or all of positions 47, 101, 106,
 CC and/or 114 of the native RNA polymerase sequence not defined in the
 CC specification. The method of the invention is utilised to identify
 CC inhibitors of the polymerase. The inhibitors serving as lead compounds
 CC for the design of potentially therapeutic compounds for the treatment of
 CC HCV. The mutant HCV NS5B RNA polymerase has improved crystallisation
 CC properties as compared to a native HCV NS5B RNA polymerase. This sequence
 CC represents a mutant RNA polymerase protein of the invention. NOTE: This
 CC sequence is not shown in the specification. It is, however, described in
 CC the specification and its sequence has been obtained from an electronic
 CC data file
 XX
 SQ Sequence 576 AA;

Query Match 96.6%; Score 2951; DB 6; Length 576;
 Best Local Similarity 96.5%; Pred. No. 7.4e-275;
 Matches 556; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

Qy 5 SMSYTTGALITPCAAESKLPINALNSLRHNLVYVTSRSASLRQKVTDFRQVL 64
 Db 1 SMSYTTGALITPCAAESKLPINALNSLRHNLVYVTSRSASLRQKVTDFRQVL 60

Qy 65 DDHYRDVLKEMKASTVKAKLSVEEACKLTPPHSAKSGFYGAKDVRSLSSRAVNHIR 124
 Db 61 DDHYRDVLKEMKASTVKAKLSVEEACKLTPPHSAKSGFYGAKDVRSLSSRAVNHIR 120

Qy 125 SVWKDLLEDTPITQTTIMAKNEVFCVQPEKGGKRPALIVFPDLGVRVCEKMAVYVVS 184
 Db 121 SVWKDLLEDTPITQTTIMAKNEVFCVQPEKGGKRPALIVFPDLGVRVCEKMAVYVVS 180

Qy 185 TLPQAVMGSSYGFQYSPKQKRVFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESII 244
 Db 181 TLPQAVMGSSYGFQYSPKQKRVFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESII 240

Qy 245 QCCDLAPEARQAIRSLTERLVVGGPMTNSKQNGYRCRASGLVLTSCGNLTTCYKAA 304
 Db 241 QCCDLAPEARQAIRSLTERLVVGGPMTNSKQNGYRCRASGLVLTSCGNLTTCYKAA 300

Qy 305 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAASLRVTEAMTRYSAAPGDPPOP 364
 Db 301 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAASLRVTEAMTRYSAAPGDPPOP 360

Qy 365 ELITSCSNVSVAHDSAGKRVYVLTDRPTVPLAARAAETARHTPVNSWLGNIIMYAPTILW 424
 Db 361 ELITSCSNVSVAHDSAGKRVYVLTDRPTVPLAARAAETARHTPVNSWLGNIIMYAPTILW 420

Qy 425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
 Db 421 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 480

Qy 485 EINRVASCLRLKGVPPPLRVVHRARSVRAKLSQGGRAAICGKYLFWNAVTKLTP 544
 Db 481 EINRVASCLRLKGVPPPLRVVHRARSVRAKLSQGGRAAICGKYLFWNAVTKLTP 540

QY 545 AASRLDLSGFWAGYSGGDIYHSLSRARPRHHHHH 580
 Db 541 AASQLDLSGFWAGYSGGDIYHSLSRARPRHHHHH 576

RESULT 3

AAO26419
 ID AAO26419 standard; protein; 576 AA.

AC AAO26419;

XX 30-JAN-2003 (first entry)

XX Mutant RNA polymerase MUT-5 protein.

DE Antiviral; hepatitis C virus; HCV; NS5B RNA polymerase.

XX Unidentified.

XX EP1256628-A2.

XX 13-NOV-2002.

PF 07-MAY-2002; 2002EP-00009387.

PR 10-MAY-2001; 2001US-0289829P.

PA (AGOU-) AGOURON PHARM INC.

PI Love RA, Yu X, Diehl W, Hickey MJ, Parge HE, Gao J, Fuhrman S;

DR WPI; 2003-031968/03.

PT Isolated nucleic acid encoding a mutant hepatitis C virus (HCV) NS5B RNA polymerase, useful for producing the mutant RNA polymerase which can be used for identifying inhibitors that can treat HCV infection.

PS Claim 9; Page; 32pp; English.

XX The invention relates to an isolated, purified nucleic acid molecule, which encodes a mutant hepatitis C virus (HCV) NS5B RNA polymerase, where a point mutation occurs at one or all of positions 47, 101, 106, and/or 114 of the native RNA polymerase sequence not defined in the specification. The method of the invention is utilised to identify inhibitors of the polymerase. The inhibitors serving as lead compounds for the design of potentially therapeutic compounds for the treatment of HCV. The mutant HCV NS5B RNA polymerase has improved crystallisation properties as compared to a native HCV NS5B RNA polymerase. This sequence represents a mutant RNA polymerase protein of the invention. NOTE: This sequence is not shown in the specification. It is, however, described in the specification and its sequence has been obtained from an electronic data file

XX Sequence 576 AA;

Query Match 96.5%; Score 2948; DB 6; Length 576;

Best Local Similarity 96.4%; Pred No. 1.4e-274;

Matches 555; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY 5 SMSYTWTCALITPCAAEESKLPINALSNLLRHNNLVYSTTSRSASLRQKKVTFDLQVL 64

Db 1 SMSYTWTCALITPCAAEESKLPINALSNLLRHNNLVYSTTSRSASLRQKKVTFDLQVL 60

QY 65 DDHYRDVLKEMKAKASTYKAKLLSVEERACKLTPPHSAKSKYGAKDVRSLSSRAVNHIR 124

Db 61 DDHYRDVLKEMKAKASTYKAKLLSVEERACKLTPPHSAKSKYGAKDVRSLSSRAVNHIR 120

QY 125 SVWKDLLEDTDPIOTTIMAXNEVFCVQPEKGRKPARLIYFPDGLGVRCMKALYDVVS 184

Db 121 SVWKDLLEDTDPIOTTIMAXNEVFCVQPEKGRKPARLIYFPDGLGVRCMKALYDVVS 180

QY 185 TLPQAVMGSSYGFQYSPQRVEFLVNTWKSKNPMGFSYDTRCFDSTVTENDIRVEESIY 244

Db 181 TLPQAVMGSSYGFQYSPQRVEFLVNTWKSKNPMGFSYDTRCFDSTVTENDIRVEESIY 240
 QY 245 OCCDLAPARQAIRSLTERLYVGGPMNTSKGQNCYRRCRAGSVLTTCGNTLTCTYKAA 304
 Db 241 OCCDLAPARQAIRSLTERLYVGGPMNTSKGQNCYRRCRAGSVLTTCGNTLTCTYKAA 300
 QY 305 AACRAAKLQDCTMLVNGDDLVI CESAGTQBDAAASLRVFTTEAMTRYSPGPPQPEYDL 364
 Db 301 AACRAAKLQDCTMLVNGDDLVI CESAGTQBDAAASLRVFTTEAMTRYSPGPPQPEYDL 360
 QY 365 ELITSCSSNVSAHDASGRVYLTTRDPTVPLARAWEETARHTPVNSLGNLIIMYAPTLM 424
 Db 361 ELITSCSSNVSAHDASGRVYLTTRDPTVPLARAWEETARHTPVNSLGNLIIMYAPTLM 420
 QY 425 ARMIIMTHFFSILLAEQLEKALDCQIYGACVSTIEPLDLPOIIRLHGLSAPLSHSYSPG 484
 Db 421 ARMIIMTHFFSILLAEQLEKALDCQIYGACVSTIEPLDLPOIIRLHGLSAPLSHSYSPG 480
 QY 485 EINRVASCLRLKGLVPPPLRVWRHRSVRARLLSOGGRAATCGKYLFWNAVTKLKTPIP 544
 Db 481 EINRVASCLRLKGLVPPPLRVWRHRSVRARLLSOGGRAATCGKYLFWNAVTKLKTPIP 540
 QY 545 AASRLDLSGFWAGYSGGDIYHSLSRARPRHHHHH 580
 Db 541 AASQLDLSGFWAGYSGGDIYHSLSRARPRHHHHH 576

RESULT 4

AAO26418

ID AAO26418 standard; protein; 576 AA.

AC AAO26418;

XX 30-JAN-2003 (first entry)

XX Mutant RNA polymerase MUT-1 protein.

XX Antiviral; hepatitis C virus; HCV; NS5B RNA polymerase.

XX Unidentified.

XX EP1256628-A2.

XX 13-NOV-2002.

XX 07-MAY-2002; 2002EP-00009387.

XX 10-MAY-2001; 2001US-0289829P.

PA (AGOU-) AGOURON PHARM INC.

PI Love RA, Yu X, Diehl W, Hickey MJ, Parge HE, Gao J, Fuhrman S;

DR WPI; 2003-031968/03.

PT Isolated nucleic acid encoding a mutant hepatitis C virus (HCV) NS5B RNA polymerase, useful for producing the mutant RNA polymerase which can be used for identifying inhibitors that can treat HCV infection.

PS Claim 8; Page; 32pp; English.

XX The invention relates to an isolated, purified nucleic acid molecule, which encodes a mutant hepatitis C virus (HCV) NS5B RNA polymerase, where a point mutation occurs at one or all of positions 47, 101, 106, and/or 114 of the native RNA polymerase sequence not defined in the specification. The method of the invention is utilised to identify inhibitors of the polymerase. The inhibitors serving as lead compounds for the design of potentially therapeutic compounds for the treatment of HCV. The mutant HCV NS5B RNA polymerase has improved crystallisation properties as compared to a native HCV NS5B RNA polymerase. This sequence represents a mutant RNA polymerase protein of the invention. NOTE: This sequence is not shown in the specification. It is, however, described in

CC the specification and its sequence has been obtained from an electronic
 CC data file
 XX
 SQ Sequence 576 AA;

Query Match 96.4%; Score 2944; DB 6; Length 576;
 Best Local Similarity 96.2%; Pred. No. 3.5e-274;
 Matches 554; Conservative 13; Mismatches 9; Indels 0; Gaps 0;
 QY 5 SMSYTWGALITPCAAEESKLPINALSNLSLRHHNLVYSTTSRSASLRQKVTFRDLQVL 64
 DB 1 SMSYTWGALITPCAAEESKLPINALSNLSLRHHNLVYSTTSRSASLRQKVTFRDLQVL 60
 QY 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYKAKVRSLSRAVNHIR 124
 DB 61 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYKAKVRSLSRAVNHIR 120
 QY 125 SVWKDLLEDVTPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 184
 DB 121 SVWKDLLEDVTPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 180
 QY 185 TLPQAVMGSSYGFQYSPQKQVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
 DB 181 TLPQAVMGSSYGFQYSPQKQVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 240
 QY 245 QCCDLAPEARQAIRSLTERLYVGGPMNTSKGQNGCYRRCRASGLVTTSCGNTLTCYLKAA 304
 DB 241 QCCDLAPEARQAIRSLTERLYVGGPMNTSKGQNGCYRRCRASGLVTTSCGNTLTCYLKAS 300
 QY 305 AACRAAKLQDCTMLVNGDDLVI CESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
 DB 301 AACRAAKLQDCTMLVNGDDLVI CESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 360
 QY 365 ELITSCSSNVSAHDASGRVYVLTTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTLW 424
 DB 361 ELITSCSSNVSAHDASGRVYVLTTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTLW 420
 QY 425 ARMLMTHTFFSILLAOEQLKALDCQIYGACYSIEPLDLPOIIRLHGLSAFSLHSYSPG 484
 DB 421 ARMLMTHTFFSILLAOEQLKALDCQIYGACYSIEPLDLPOIIRLHGLSAFSLHSYSPG 480
 QY 485 EINRVASCLRLKGLVPPPLRVWRHRARSVRARLLSQGGRAAICGKYLFWNAVRTKLTPI 544
 DB 481 EINRVASCLRLKGLVPPPLRVWRHRARSVRARLLSQGGRAAICGKYLFWNAVRTKLTPI 540
 QY 545 AASRLDLSCGFVAGYSGGDIYHSLSRARPRHHHHH 580
 DB 541 AASRLDLSCGFVAGYSGGDIYHSLSRARPRHHHHH 576

RESULT 5
 AAB60131
 ID AAB60131 standard; protein; 579 AA.
 XX AAB60131;
 AC
 XX
 DT 03-APR-2001 (first entry)
 XX
 DE Hepatitis C virus NS5B polymerase related protein.
 XX
 KW Hepatitis C virus; HCV; NS5B polymerase; protein coordinate data;
 KW cirrhosis; hepatoma; inhibitor.
 XX
 OS Hepatitis C virus.
 XX
 PN EP1065213-A2.
 XX
 PD 03-JAN-2001.
 XX
 PF 30-JUN-2000; 2000EP-00113955.
 XX
 PR 02-JUL-1999; 99JP-00188630.
 PR 07-JUL-1999; 99JP-00192488.

XX (NISB) JAPAN TOBACCO INC.
 PA Ago H, Miyano M, Adachi T;
 FI WPI; 2001-104834/12.
 DR New polypeptide, derived from hepatitis C virus (HCV) polymerase NS5B,
 XX having a HCV polymerase activity, useful for crystal structure analysis
 PT and for rational identification of HCV polymerase inhibitors.
 XX Disclosure; Page 279-282; 295pp; English.
 PS
 CC The present invention provides the hepatitis C virus (HCV) NS5B
 CC polymerase protein and its structure. It also provides methods for
 CC identifying inhibitors of the protein which can be used in the treatment
 CC of HCV infection. HCV can lead to cirrhosis and hepatoma in its chronic
 CC form
 XX
 SQ Sequence 579 AA;

Query Match 96.3%; Score 2942; DB 4; Length 579;
 Best Local Similarity 96.0%; Pred. No. 5.5e-274;
 Matches 556; Conservative 12; Mismatches 9; Indels 2; Gaps 1;
 QY 4 MSMSYTWGALITPCAAEESKLPINALSNLSLRHHNLVYSTTSRSASLRQKVTFRDLQV 63
 DB 1 MSMSYTWGALITPCAAEESKLPINALSNLSLRHHNLVYSTTSRSASLRQKVTFRDLQV 60
 QY 64 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYKAKVRSLSRAVNHIR 123
 DB 61 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYKAKVRSLSRAVNHIR 120
 QY 124 RSVWKDLLEDVTPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVV 183
 DB 121 RSVWKDLLEDVTPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVV 180
 QY 184 STLPQAVMGSSYGFQYSPQKQVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESI 243
 DB 181 STLPQAVMGSSYGFQYSPQKQVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESI 240
 QY 244 YQCCDLAPEARQAIRSLTERLYVGGPMNTSKGQNGCYRRCRASGLVTTSCGNTLTCYLKA 303
 DB 241 YQCCDLAPEARQAIRSLTERLYVGGPMNTSKGQNGCYRRCRASGLVTTSCGNTLTCYLKA 300
 QY 304 AACRAAKLQDCTMLVNGDDLVI CESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYD 363
 DB 301 AACRAAKLQDCTMLVNGDDLVI CESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYD 360
 QY 364 LELITSCSSNVSAHDASGRVYVLTTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTL 423
 DB 361 LELITSCSSNVSAHDASGRVYVLTTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTL 420
 QY 424 WARMILMTHTFFSILLAOEQLKALDCQIYGACYSIEPLDLPOIIRLHGLSAFSLHSYSP 483
 DB 421 WARMILMTHTFFSILLAOEQLKALDCQIYGACYSIEPLDLPOIIRLHGLSAFSLHSYSP 480
 QY 484 GEINRVASCLRLKGLVPPPLRVWRHRARSVRARLLSQGGRAAICGKYLFWNAVRTKLTPI 543
 DB 481 GEINRVASCLRLKGLVPPPLRVWRHRARSVRARLLSQGGRAAICGKYLFWNAVRTKLTPI 540
 QY 544 PAASRLDLSCGFVAGYSGGDIYHSLSRARPR--HHHHH 580
 DB 541 PAASRLDLSCGFVAGYSGGDIYHSLSRARPRGGSHHHH 579

RESULT 6
 AAG79558
 ID AAG79558 standard; protein; 578 AA.
 XX AAG79558;
 AC
 XX
 DT 09-DEC-2002 (first entry)

XX HCV decrease-affinity NS5B polymerase, NS5Bdelta21C-HT.
 XX Enzyme; hepatitis C virus; HCV; decreased affinity; NS5B; polymerase;
 KW inhibitor; RNA-dependent RNA polymerase.
 XX Hepatitis C virus.
 OS WO200270739-A2.
 XX 12-SEP-2002.
 XX 06-MAR-2002; 2002WO-CA000323.
 XX 08-MAR-2001; 2001US-0274374P.
 XX (BOEH) BOEHRINGER INGELHEIM CANADA LTD.
 XX Kukolj G, Mckercher G;
 XX WPI; 2002-698759/75.
 XX Identifying inhibitors HCV NS5B RNA-dependent RNA polymerase, comprises
 PT incubating an HCV NS5B polymerase having decreased affinity for the
 PT primer template relative to the native polymerase, with a potential
 PT inhibitor.
 XX Claim 9; Page 47-48; 52pp; English.
 XX The sequences given in AAG79556-61 are hepatitis C virus (HCV) decreased
 CC affinity NS5B polymerases. These enzymes are used in the method of the
 CC invention for identifying a potential inhibitor of the binding between a
 CC HCV NS5B RNA-dependent RNA polymerase and an appropriate primer-template.
 CC The method comprises incubating the HCV NS5B polymerase with the primer-
 CC template in the presence and absence of a potential inhibitor. The HCV
 CC NS5B polymerase has a decreased affinity for the primer-template relative
 CC to that of native HCV NS5B RNA-dependent RNA polymerase. The method is
 CC useful for identifying a potential inhibitor of the binding between a HCV
 CC NS5B RNA-dependent RNA polymerase and a primer-template. Use of
 CC polymerase constructs having a lower affinity towards the primer-template
 CC than that of native NS5B polymerase is particularly useful for
 CC identifying potential inhibitors in screening large libraries of
 CC compounds. The new method reduces the difficulties and disadvantages of
 CC prior art. The present method provides an assay that is easy to perform
 CC on large libraries of compounds, and has improved sensitivity for
 CC detecting inhibitors that would not be identified as such using native
 CC NS5B polymerase. This protein represents a soluble form of mature HCV
 CC NS5B which lacks the C-terminal 21 amino acids and has an C-terminal hexa
 CC -histidine tag
 XX Sequence 578 AA;
 SQ
 Query Match 96.1%; Score 2936; DB 5; Length 578;
 Best Local Similarity 95.5%; Pred. No. 2.1e-273;
 Matches 552; Conservative 18; Mismatches 6; Indels 2; Gaps 1;
 QY 5 SMSYTWTCALITPCAABESKLPINALNSLRHNLVSTTSRSASLRQKKVTFDRQLVL 64
 DB 1 SMSYTWTCALITPCAABESQIPINALNSLRHNLVSTTSRSALRQKKVTFDRQLVL 60
 QY 65 DDHYRDVLKEMKAKASTYKAKLLSYEEACKLTPPHSAKSKFGYGAQVRSLSRAVNHIR 124
 DB 61 DDHYRDVLKEMKAKASTYKAKLLSYEEACKLTPPHSAKSKFGYGAQVRSLSRAVDHIR 120
 QY 125 SWKDLLEDTPDTPIOTTTIMAKNEVFCVQPEKGRKPARLIIVPDLGVRCMKALYDVWS 184
 DB 121 SWKDLLEDTPDTPIOTTTIMAKNEVFCVQPEKGRKPARLIIVPDLGVRCMKALYDVWS 180
 QY 185 TLPAVMGSSYGFQYSPKQRYVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVESIY 244
 DB 181 TLPAVMGSSYGFQYSPKQRYVEFLVNTWKAKCPMGFSYDTRCFDSTVTESDIRVESIY 240
 QY 245 QCCDLAPARQAIRSLTERLYVGGPMTNSKGQNGCYRRRCRAGSVLTTCGNTLTCLYKAA 304

DB 241 QCCDLAPARQAIRSLTERLYVGGPMTNSKGQNGCYRRRCRAGSVLTTCGNTLTCLYKAS 300
 QY 305 AACRAAKLQDCTMLVNGDDLAVVICESAGTQSDAASLRVFTAMTRYSAAPPDPPEYDL 364
 DB 301 AACRAAKLQDCTMLVNGDDLAVVICESAGTQSDAASLRVFTAMTRYSAAPPDPPEYDL 360
 QY 365 ELITSCSSNVSAHDASGKRVYVLTTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTILW 424
 DB 361 ELITSCSSNVSAHDASGKRVYVLTTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTILW 420
 QY 425 ARMLTMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAPLSHSYSPG 484
 DB 421 ARMLTMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAPLSHSYSPG 480
 QY 485 EINRVASCLRLKLGVPPLRVWRHRSVRKALLSQGGRRAICGKYLFWNAVRTKLTPTIP 544
 DB 481 EINRVASCLRLKLGVPPLRVWRHRSVRKALLSQGGRRAICGKYLFWNAVRTKLTPTIP 540
 QY 545 AASRLDLGSGWFGVAGYSGGDIYHSLSRAPR--HHHHHH 580
 DB 541 AASRLDLGSGWFGVAGYSGGDIYHSLSRAPRLEHHHHH 578
 RESULT 7
 ID ABP71662 standard; protein; 578 AA.
 XX ABP71662;
 AC ABP71662;
 XX 29-MAY-2003 (first entry)
 DT HCV NS5Bdelta21-His protein sequence.
 DE HCV NS5Bdelta21-His protein sequence.
 KW HCV; polymerase; virucide; hepatotropic; antiviral; enzyme; NS5B.
 OS Hepatitis C virus.
 XX WO2003014377-A2.
 XX 20-FEB-2003.
 XX 05-AUG-2002; 2002WO-CA001214.
 XX 07-AUG-2001; 2001US-0310272P.
 XX (BOEH) BOEHRINGER INGELHEIM CANADA LTD.
 XX Kukolj G, Beaulieu PL, Mckercher G;
 XX WPI; 2003-300535/29.
 XX Identifying compounds that bind to hepatitis C virus polymerase, useful
 PT as antiviral agents, from ability to displace bound probe from the
 PT polymerase.
 XX Example 2; Page 117-119; 125pp; English.
 XX The invention relates to a method for identifying compounds that bind to
 CC hepatitis C virus (HCV) polymerase. The method involves (a) treating HCV
 CC polymerase, or its analogue, with a probe that binds to the polymerase to
 CC form a complex but is displaceable by an inhibitor; (b) measuring a
 CC signal emitted from the probe in the complex to establish a baseline
 CC value; (c) incubating the complex with test compound; and (d) measuring
 CC the signal from the complex and comparing it with the baseline. Any
 CC modulation indicates that the test compound binds to HCV polymerase. The
 CC method is used to identify (also to determine potency) of HCV polymerase
 CC inhibitors, potentially useful as antiviral agents. The present sequence
 CC represents a HCV NS5Bdelta21-His polymerase
 XX Sequence 578 AA;
 SQ
 Query Match 96.1%; Score 2936; DB 6; Length 578;

Best Local Similarity		95.5%;	Pred. No. 2.1e-273;	
Matches		552;	Conservative	18; Mismatches
		6;	Indels	2; Gaps
QY	5	SMSTVTGALITPCAAEESKLPINALNSLIRHNLVYSTTSRASLIRKQKVTFRDLQVL	64	
DB	1	SMSTVTGALITPCAAEESQLPINALNSLVRHNMVYSTTSRAALRKQKVTFRDLQVL	60	
QY	65	DDHYRDVLKEMKAKASTVYAKLLSVEEACKLTPPHSAKSKFGYGAQDVRSLSRAVNHIR	124	
DB	61	DDHYRDVLKEMKAKASTVYAKLLSVEEACKLTPPHSAKSKFGYGAQDVRLSSKAVDHIR	120	
QY	125	SVWKDLLEDTDTPITTIMAKNEVFCVQPEKGGKRPARLIVFPDLGVRVRCRMALYDVVS	184	
DB	121	SVWKDLLEDTETPIDTIMAKNEVFCVQPEKGGKRPARLIVFPDLGVRVRCRMALYDVVS	180	
QY	185	TLPQAVNGSSYGFQYSPKQVPELVNTWKAKKCPMGFSYDFTRCFDSTVTENDIRVEESIY	244	
DB	181	TLPQAVNGSSYGFQYSPKQVPELVNTWKAKKCPMGFSYDFTRCFDSTVTESDIRVEESIY	240	
QY	245	QCDDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNLTLCYLKAA	304	
DB	241	QCDDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNLTLCYLKAS	300	
QY	305	AACRAAKLOCTMLVNGDDLVIICESAGTQBDAAASLRVFTAMTRYSAPOGDPQPEYDL	364	
DB	301	AACRAAKLOCTMLVNGDDLVIICESAGTQBDAAANLRVFTAMTRYSAPOGDLQPEYDL	360	
QY	365	ELITSSCSNVSVAHDASGKRVYILTRDPTVPLARAAMETABHTVPVNSLGNIMVAPTLW	424	
DB	361	ELITSSCSNVSVAHDASGKRVYILTRDPTVPLARAAMETABHTVPVNSLGNIMVAPTLW	420	
QY	425	ARMILMTHFFSILLIAQLEKALDCQIYGACYSIEPLDLPOIIERLHGLSAFSLHSYSPG	484	
DB	421	ARMVLMTHFFSILLIAQLEKALDCQIYGACYSIEPLDLPOIIERLHGLSAFSLHSYSPG	480	
QY	485	EINRVASCLRLGVPPPLVRVHRARSVRAKLLSQGGRAAICGKYLEFNWAVRTKLTPI	544	
DB	481	EINRVASCLRLGVPPPLVRVHRARSVRAKLLSQGGRAATCGKYLEFNWAVRTKLTPI	540	
QY	545	AASRLDLSGDFVAGYSGDDIYHLSRARPR--HHHHH	580	
DB	541	AASRLDLSGDFVAGYSGDDIYHLSRARPRLEHHHHH	578	
RESULT 8				
ABG72214				
ID	ABG72214 standard; protein; 576 AA.			
XX	ABG72214;			
XX				
XX				
XX				
XX	17-FEB-2003 (first entry)			
XX				
DE	HCV mutant NS5B polymerase polypeptide, His-NS5BdelC721.			
DE				
XX				
KW	Crystallisation; crystal structure; hepatitis C virus; HCV;			
KW	NS5B polymerase; NS5B RNA-dependent RNA polymerase; X-ray diffraction;			
KW	atomic coordinate determination; drug screening; mutant; mutein;			
KW	protein co-ordinate data; enzyme; His-NS5BdelC721.			
XX				
OS	Hepatitis C virus; BK isolate.			
OS	Synthetic.			
XX				
XX	Key			
FT	Location/Qualifiers			
FT	Misc-difference 335			
FT	/label= Thr, Val			
FT	/note= "Preferred amino acid at this position is valine."			
FT	Wild type residue at this position is threonine"			
FT	Misc-difference 344			
FT	/label= Val, Ala			
FT	/note= "Preferred amino acid at this position is alanine."			
FT	Wild type residue at this position is valine"			
FT	Misc-difference 550			
FT	/label= Arg, Gln			

FT	/note= "Preferred amino acid at this position is
FT	glutamine. Wild type residue at this position is
FT	arginine"
PN	US6434489-B1.
XX	13-AUG-2002.
PD	03-APR-2000; 2000US-00541990.
XX	03-APR-2000; 2000US-00541990.
XX	(SCHE) SCHERING CORP.
XX	Lesburg CA, Cable M, Hong Z, Mannarino AF, Weber PC;
XX	WPI; 2003-110131/10.
XX	New crystalline Hepatitis C Virus NS5B polypeptide composition useful for
PT	determining the atomic coordinates of the polypeptide and drug screening.
PT	Claim 1; Col 75-80; 44pp; English.
XX	The present invention relates to the purification, crystallisation, and
CC	structure of hepatitis C virus (HCV) NS5B RNA-dependent RNA polymerase.
CC	The composition containing the crystals of HCV NS5B polymerase
CC	polypeptide diffracts x-rays for determination of the atomic coordinates
CC	of the polypeptide to greater than 4.0 Angstroms resolution. The crystal
CC	is useful for determination of the atomic coordinates of the HCV NS5B
CC	polypeptide, and is useful in drug screening methods. The present
CC	sequence represents a mutant of the published sequence (Genbank Accession
CC	No. 130458) for the HCV NS5B polypeptide, His-NS5BdelCT21. This mutant
CC	contains 3 amino acid substitutions and a polyhistidine tag at the N-
CC	terminus. The C-terminal 21 amino acids present in the wild type are also
CC	deleted in His-NS5BdelCT21
XX	Sequence 576 AA:
XX	SO

Query Match	94.2%	Score 2879;	DB 6;	Length 576;
Best Local Similarity	96.0%;	Pred. No. 6.5e-268;		
Matches 545; Conservative	11;	Missmatches 12;	Indels 0;	Gaps 0;

QY	7	SYTWTGALITPCAAEBSKLPINALNSLLRRHNLVYTTSRASLRQKKVTFDRQLQVLDD	66
DB	9	SYTWTGALITPCAAEBSKLPINALNSLLRRHNNVYATTSRAGLRQKKVTFDRQLQVLDD	68
QY	67	HYRDVLKEMKAKASTVYKALLSVEEACKLTPPHSAKSKFGYGAQKDVLSRLSRVNHHSV	126
DB	69	HYRDVLKEMKAKASTVYKALLSVEEACKLTPPHSAKSKFGYGAQKDVRLSSKAVNHHISV	128
QY	127	WKDLLEDTPDIOTITMAKNNEVFCVQPEKGGRKPARLIIVPDLGVRVCEKMALEDVYVSTL	186
DB	129	WKDLLEDTVPIDITITMAKNNEVFCVQPEKGGRKPARLIIVPDLGVRVCEKMALEDVYVSTL	188
QY	187	PQAVMGSSYGFQYSPQORVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIYQC	246
DB	189	PQVVMGSSYGFQYSPQORVEFLVNTWKSXKMPGFSYDTRCFDSTVTENDIRVEESIYQC	248
QY	247	CDLAPEARQAIKSLTERLYVGGPMTNSKGQNGCYRRCRASGLVTTSCGNTLITCYLKAAA	306
DB	249	CDLAPEARQAIKSLTERLYVGGPMTNSKGQNGCYRRCRASGLVTTSCGNTLITCYLKASNA	308
QY	307	CRAAKLQDCTMLVNGDDLVIICBSAGTQEDAAASLRVFTEAMTRVYSPAPGDPQPEYDLEL	366
DB	309	CRAAKLQDCTMLVNGDDLVIICBSAGTQEDAAASLRXFTEAMTRVYSPAPGDPQPEYDLEL	368
QY	367	ITSCSSNVSVAHADSGKRVYLLTRDPTVPLARAAMETARHTPPVNSWLGNIIMYAPTLWAR	426
DB	369	ITSCSSNVSVAHADSGKRVYLLTRDPTVPLARAAMETARHTPPVNSWLGNIIMYAPTLWAR	428
QY	427	MILMTHFFSILLAQEOLEKALDCQIYGCACYSIEPLDLPQIIRLHGLHSAFSLHSYSGPEI	486
DB	429	MILMTHFFSILLAQEOLEKALDCQIYGCACYSIEPLDLPQIIRLHGLHSAFSLHSYSGPEI	488

XX Hyun US, Kim JM, Lee MG, Noh GY;
 XX WPI; 2003-253755/25.
 DR N-PSDB; ADM69006.
 XX Recombinant hepatitis C virus ns5b protein, and preparation process and
 PT use thereof.
 XX Claim 3; SEQ ID NO 8; 19pp; Korean.
 XX The invention relates to a novel recombinant Hepatitis C virus (HCV) NS5B
 CC protein which is derived from a HCV patient's serum and has excellent RNA
 CC -dependent RNA polymerase (RdRp) activity. The polypeptide of the
 CC invention may be useful as a HCV growth inhibitor. The current sequence
 CC is that of the HCV NS5B RNA-dependent RNA polymerase (RdRp) protein of
 CC the invention.
 XX Sequence 567 AA;
 SQ
 Query Match 93.2%; Score 2846; DB 7; Length 567;
 Best Local Similarity 94.7%; Pred. No. 9.6e-265;
 Matches 537; Conservative 17; Mismatches 13; Indels 0; Gaps 0;
 QY 5 SMSYTWTCALITPCAAEESKLPINALNSLRLHNLVYSTTSRSASLQKQKVTFRDLQVL 64
 Db 1 SMSYSWTGALITPCAAEESKLPINPLNSLRLHNLVYATTSSAGLQKQKVTFRDLQVL 60
 QY 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAADVRSLSRAVNHIR 124
 Db 61 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAADVRSLSRAVNHIR 120
 QY 125 SVWKDLLEDTTPIOTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRCVKMALYDVVS 184
 Db 121 SVWKDLLEDTTPIOTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRCVKMALYDVVS 180
 QY 185 TLPQAVMGSSYGFQYSPQRVFEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
 Db 181 TLPQAVMGSSYGFQYSPQRVFEFLVNTWKAKKCPMGFSYDTRCFDSTVTESDIRVEESIY 240
 QY 245 QCCDLAPEARQAIRSLTERLYVGGPMNTSKGNCYRCRASGLVLTSCGNTLCYLKAA 304
 Db 241 QCCDLAPAKQAIRSLTERLYVGGPLTNSKGNCYRCRASGLVLTSCGNTLCYLKAS 300
 QY 305 AACRAAKLODCTMLVNGDVLVVCESACTQEDAAASLRVTEAMTRYSAPPDGPQPEYDL 364
 Db 301 AACRAAKLODCTMLVNGDVLVVCESACTQEDAAASLRVTEAMTRYSAPPDGPQPEYDL 360
 QY 365 ELITSCSNVSAHDASGRVYVLTTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTLW 424
 Db 361 ELITSCSNVSAHDASGRVYVLTTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTLW 420
 QY 425 ARMILMTHFFSILLAOEALDQCIYGCYSIEPLDLPQIIRLHGLSAFSLHSYSPG 484
 Db 421 ARMILMTHFFSILLAOEALDQCIYGCYSIEPLDLPQIIRLHGLSAFSLHSYSPG 480
 QY 485 EINRVASCLRLKGVPLRWRHRAVSRAKLLSQGGRAAICGKYLFWAVRTKLTPIIP 544
 Db 481 EINRVASCLRLKGVPLRWRHRAVSRAKLLSQGGRAATCGKYLFWAVRTKLTPIIP 540
 QY 545 AASRLDLSGWFVAGYSGGDIYVSLSRA 571
 Db 541 AASRLDLSGWFVAGYSGGDIYVSLSRA 567

RESULT 11

AAG79559

ID AAG79559 standard; protein; 545 AA.

XX AAG79559;

AC AAG79559;

XX 09-DEC-2002 (first entry)

DT

XX

DE HCV decrease-affinity NS5B polymerase, NS5Bdelta57-HT.
 XX Enzyme; hepatitis C virus; HCV; decreased affinity; NS5B; polymerase;
 KW inhibitor; RNA-dependent RNA polymerase.
 XX Hepatitis C virus.
 OS WO200270739-A2.
 PN 12-SEP-2002.
 PD 06-MAR-2002; 2002WO-CA000323.
 XX 08-MAR-2001; 2001US-0274374P.
 PR (BOEH) BOEHRINGER INGELHEIM CANADA LTD.
 XX Kukolj G, Mckercher G;
 FI WPI; 2002-698759/75.
 XX Identifying inhibitors HCV NS5B RNA-dependent RNA polymerase, comprises
 DR incubating an HCV NS5B polymerase having decreased affinity for the
 XX primer template relative to the native polymerase, with a potential
 XX inhibitor.
 PS Claim 10; Page 48-49; 52pp; English.
 XX The sequences given in AAG79556-61 are hepatitis C virus (HCV) decreased
 CC affinity NS5B polymerases. These enzymes are used in the method of the
 CC invention for identifying a potential inhibitor of the binding between a
 CC HCV NS5B RNA-dependent RNA polymerase and an appropriate primer-template.
 CC The method comprises incubating the HCV NS5B polymerase with the primer-
 CC template in the presence and absence of a potential inhibitor. The HCV
 CC NS5B polymerase has a decreased affinity for the primer-template relative
 CC to that of native HCV NS5B RNA-dependent RNA polymerase. The method is
 CC useful for identifying a potential inhibitor of the binding between a HCV
 CC NS5B RNA-dependent RNA polymerase and a primer-template. Use of
 CC polymerase constructs having a lower affinity towards the primer-template
 CC than that of native NS5B polymerase is particularly useful for
 CC identifying potential inhibitors in screening large libraries of
 CC compounds. The new method reduces the difficulties and disadvantages of
 CC prior art. The present method provides an assay that is easy to perform
 CC on large libraries of compounds, and has improved sensitivity for
 CC detecting inhibitors that would not be identified as such using native
 CC NS5B polymerase. This protein represents a soluble form of mature HCV
 CC NS5B which lacks the C-terminal 57 amino acids and has an C-terminal hexa
 CC -histidine tag
 XX Sequence 545 AA;
 SQ
 Query Match 89.5%; Score 2734.5; DB 5; Length 545;
 Best Local Similarity 90.1%; Pred. No. 5.1e-254;
 Matches 519; Conservative 19; Mismatches 7; Indels 31; Gaps 2;
 QY 5 SMSYTWTCALITPCAAEESKLPINALNSLRLHNLVYSTTSRSASLQKQKVTFRDLQVL 64
 Db 1 SMSYTWTCALITPCAAEESKLPINALNSLRLHNLVYSTTSRSASLQKQKVTFRDLQVL 60
 QY 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAADVRSLSRAVNHIR 124
 Db 61 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAADVRSLSRAVNHIR 120
 QY 125 SVWKDLLEDTTPIOTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRCVKMALYDVVS 184
 Db 121 SVWKDLLEDTTPIOTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRCVKMALYDVVS 180
 QY 185 TLPQAVMGSSYGFQYSPQRVFEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
 Db 181 TLPQAVMGSSYGFQYSPQRVFEFLVNTWKAKKCPMGFSYDTRCFDSTVTESDIRVEESIY 240
 QY 245 QCCDLAPEARQAIRSLTERLYVGGPMNTSKGNCYRCRASGLVLTSCGNTLCYLKAA 304
 XX

Db 241 QCCDLAPARQAISLTERLYIGGPIUTNSKGQNGCYRRCRASGVLTTSCGNTLTCTVLKAS 300
 QY 305 AACRAAKLQDCTMLVNGDDLVIWICSSAGTQEDAAASLRVFTTEAMTRYSAAPGPPQPEYDL 364
 Db 301 AACRAAKLQDCTMLVNGDDLVIWICSSAGTQEDAAANLRVFTTEAMTRYSAAPGDLPOPEYDL 360
 QY 365 ELITSCSSNVVAHDASGRVYVLTDRPTVPLARAAMETARHTPVNSWLGNIIMYAPTILW 424
 Db 361 ELITSCSSNVVAHDASGRVYVLTDRPTTPLARAAMETARHTPINSWLGNIIMYAPTILW 420
 QY 425 ARWMLMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSFAFSLHSYSPG 484
 Db 421 ARWMLMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSFAFSLHSYSPG 480
 QY 485 EINRVASCLRLGVPPLRVWRHARSVRAKLSQGGRAAICGKYLFWNAVRTKLTLPTIP 544
 Db 481 EINRVASCLRLGVPPLRVWRHARSVRAKLSQGGRAATCGKYLFWNAVRTKL----- 534
 QY 545 AASRLDLSGFWAGYSGGDIYHLSLRARPRHHHHH 580
 Db 535 -AAALE-----HHHHH 545

RESULT 12

AAR58592

ID AAR58592 standard; protein; 462 AA.

AC AAR58592;

XX

DT 09-MAY-1995 (first entry)

XX

DE Hepatitis C virus 8003-9388 fragment antigen.

XX

KW Non-A non-B hepatitis virus antigens; NANBH; hepatitis C virus.

XX

OS Hepatitis C virus.

XX

PN JP06225770-A.

XX

PD 16-AUG-1994.

XX

PF 08-JUL-1993; 93JP-00193104.

XX

PR 10-JUL-1992; 92JP-00207391.

XX

PA (TOKR-) ZH TOKYOTO RINSHO IGAKU SOGO KENKYUSHO.

XX

PA (SANW) SANWA KAGAKU KENKYUSHO CO.

XX

PA (TOFU) TONEN CORP.

XX

PA (KOKU-) KOKUSAI SHIYAKU KK.

XX

XX WPI; 1994-298800/37.

DR

DR N-PSDB; AAQ70542.

XX

XX A nucleic acid fragment coding Non-A Non-B Hepatitis virus antigens - for

PT diagnosis of NANBH and detection of HCV.

XX

PS Claim 1; Page 16; 22pp; Japanese.

XX

XX AAQ70542 is a fragment of hepatitis C virus (HCV) or non-A non-B hepatitis

CC virus (NANBH) non-structural coding region five, encompassing base pairs

CC 8003-9388. It codes for AAR58592 an antigen to a non-structural region of

CC the HCV virus, which can be used in the diagnosis of NANBH patients and

CC the detection of HCV carriers

XX

XX Sequence 462 AA;

XX

XX Query Match 74.6%; Score 2279; DB 2; Length 462;

XX Best Local Similarity 93.0%; Pred. No. 3.2e-210;

XX Matches 426; Conservative 18; Mismatches 14; Indels 0; Gaps 0;

XX

QY 117 SRVNHIRSVKMDLLEDDTTPIQTITIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEK 176

Db 1 SKAVNHIRSVKMDLLEDDNVTPTDITWAKSEVFCVQPEKGRKPARLIVFPDLGVRVCEK 60

QY 177 MALYDVVSTLPOAVMGSSYGFQYSPKQRFVFLVNTWKAKKCPMGFSYDTRCFDSTVTEND 236
 Db 61 MALYDVVSTLPOAVMGSSYGLQYSPQRFVFLVNAWKSKCPMGFSYDNRCDSTVTESD 120
 QY 237 IRVBSIYQCCDLAPARQAISLTERLYIGGPIUTNSKGQNGCYRRCRASGVLTTSCGNT 296
 Db 121 IRVBSIYQCCDLAPARQAISLTERLYIGGPIUTNSKGQNGCYRRCRASGVLTTSCGNT 180
 QY 297 LTCYLKAAACRAAKLQDCTMLVNGDDLVIWICSSAGTQEDAAASLRVFTTEAMTRYSAAPGD 356
 Db 181 LTCYLKAAACRAAKLQDCTMLVNGDDLVIWICSSAGTQEDAAASLRVFTTEAMTRYSAAPGD 240
 QY 357 PPOPEYDLELITSCSSNVSVAHDSAGKEVYVLTDRPTVPLARAAMETARHTPVNSWLGNI 416
 Db 241 PPRPEYDLELITSCSSNVSVAHDSAGKEVYVLTDRPTVPLARAAMETARHTPVNSWLGNI 300
 QY 417 IMYAPTILWARMILMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAF 476
 Db 301 IMYAPTILWARMILMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAF 360
 QY 477 SLHSYSPGEINRVASCLRLGVPPLRVWRHARSVRAKLSQGGRAAICGKYLFWNAVRT 536
 Db 361 SLHSYSPGEINRVASCLRLGVPPLRVWRHARSVRAKLSQGGRAATCGKYLFWNAVRT 420
 QY 537 KLTLPTIPAAASRLDLSGFWAGYSGGDIYHLSLRARPR 574
 Db 421 KLTLPTIPAAASRLDLSGFWAGYSGGDIYHLSLRARPR 458

RESULT 13

ADL17783

ID ADL17783 standard; protein; 462 AA.

XX

AC ADL17783;

XX

DT 06-MAY-2004 (first entry)

XX

DE Hepatitis C virus protein (clone C6-82).

XX

KW non-A non-B type hepatitis virus antigen;

KW hepatitis C virus antibody detection; clone C260-1.

XX

OS Hepatitis C virus.

XX

XX JP2004000150-A.

XX

XX 08-JAN-2004.

XX

XX 24-FEB-2003; 2003JP-00046379.

XX

XX 10-JUL-1992; 92JP-00207391.

XX

XX 08-JUL-1993; 93JP-00193104.

XX

XX (KOKU-) KOKUSAI SHIYAKU KK.

XX

XX WPI; 2004-085213/09.

XX

XX N-PSDB; ADL17778.

XX

XX Novel non-A non-B type hepatitis virus nucleic acid fragment useful for

PT diagnosing non-A non-B-hepatitis patient and non-A non-B type hepatitis

PT virus carrier.

XX

XX Disclosure; Page 32-33; 35pp; Japanese.

XX

XX This invention relates to a novel isolated nucleic acid fragment encoding

CC a non-A non-B type hepatitis virus antigen protein sequence.

CC Specifically, it refers to a method for diagnosing non-A non-B type

CC hepatitis patients, as well as for detecting those people who are

CC carriers. The present invention describes a diagnosis method that uses a

CC hepatitis C virus antibody detection reagent and subsequent PCR

CC amplification to identify a clone C260-1 that encodes the non-A non-B

CC type hepatitis virus antigen sequence. This polypeptide sequence is the

CC non-A-non-B type hepatitis virus protein fragment of the invention. NOTE:
 CC This sequence is given in the sequence listing as an embedded protien and
 CC is not further referred to in the specification.

XX SQ Sequence 462 AA;
 Query Match 74.6%; Score 2279; DB 8; Length 462;
 Best Local Similarity 93.0%; Pred. No. 3.2e-210;
 Matches 426; Conservative 18; Mismatches 14; Indels 0; Gaps 0;
 QY 117 SRVNHRSVWKLLEDVDTPIQTIMAKNEVFCVQPEKGRKPARLIVFDLGVVCEK 176
 Db 1 SKVNHRSVWKLLEDVDTPIQTIMAKNEVFCVQPEKGRKPARLIVFDLGVVCEK 60
 QY 177 MALYDVVSTLQAVMGSSYGQYSPKORVEFLVNTWKAKKCPMGFSYDTRCFDSTVTND 236
 Db 61 MALYDVVSTLQAVMGSSYGQYSPKORVEFLVNTWKAKKCPMGFSYDTRCFDSTVTND 120
 QY 237 IRVESIYQCCDLAPPEARQAIRSLTERLYVGGPMTNSKGQYRCRASGLVLTSCGNT 296
 Db 121 IRVESIYQCCDLAPPEARQAIRSLTERLYVGGPMTNSKGQYRCRASGLVLTSCGNT 180
 QY 297 LTCYLKAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAAASLRFVTEATRYSAAPGD 356
 Db 181 LTCYLKAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAAASLRFVTEATRYSAAPGD 240
 QY 357 PPOEYDELITSCSSNVSVADHAGSKRVYVLTTRDPTVPLARAATWETARHTFPVNSWLGNI 416
 Db 241 PPRPEYDELITSCSSNVSVADHAGSKRVYVLTTRDPTVPLARAATWETARHTFPVNSWLGNI 300
 QY 417 IMYAPTLWARMLMTHFFSILLAOBLEKALDCQIYGACYSIEPLDLPOIIRLHGLSAP 476
 Db 301 IMYAPTLWARMLMTHFFSILLAOBLEKALDCQIYGACYSIEPLDLPOIIRLHGLSAP 360
 QY 477 SLHYSYSGEINRVASCLKGLGVPLVRVWRHARSVRAKLLSQGGRAATCGKYLFWNAVRT 536
 Db 361 SLHYSYSGEINRVASCLKGLGVPLVRVWRHARSVRAKLLSQGGRAATCGKYLFWNAVRT 420
 QY 537 KLKLTPIPAASRLDLSGWFGVAGYSGGDIYHLSRARPR 574
 Db 421 KLKLTPIPAASRLDLSGWFGVAGYSGGDIYHLSRARPR 458

RESULT 14
 AAR67630
 ID AAR67630 standard; protein; 547 AA.
 XX AC AAR67630;
 XX DT 27-AUG-2003 (revised)
 XX DT 25-MAR-2003 (revised)
 XX DT 18-AUG-1995 (first entry)
 XX DE Non-A Non-B hepatitis virus non-structural protein.
 XX KW Non-A Non-B hepatitis virus; structural region; cDNA to genomic RNA;
 XX KW detection; reagent; anti-Non-A Non-B hepatitis virus antibody; vaccine;
 XX KW antigen; epitope; diagnosis.
 XX OS Non-A.
 XX OS non-B hepatitis virus.

XX Key Location/Qualifiers
 XX FT Peptide 52.72
 XX FT /note= "putative epitope site (see AAR67641)"

XX EP628572-A2.
 XX PD 14-DEC-1994.
 XX XX 27-MAY-1994; 94EP-00108256.
 XX PR 28-MAY-1993; 93JP-00126709.

PR 02-MAR-1994; 94JP-00032201.
 XX (EISA) EISAI CO LTD.
 XX Aoyama M, Obara T, Tohmatsu J, Sawada T, Hosoda T, Iwasaki Y;
 FI Arima T;
 XX WPI; 1995-015655/03.
 XX N-PSDB; AAQ75819.
 PT New non-A non-B hepatitis virus sub-type - used to develop prods. for
 PT detection, diagnosis, prevention and treatment of non-A non-B hepatitis.
 XX Claim 1; Page 33-37; 59pp; English.
 XX AAR67630 is the structural protein of the Non-A Non-B (NANB) hepatitis
 CC virus encoded by a partial cDNA to genomic RNA sequence. Regions of the
 CC polypeptide were studied for suitability as an epitope. The novel
 CC epitopes are effective in the diagnosis of NANB hepatitis. The nucleotide
 CC sequences (see AAQ75817-19) were isolated from the plasma of donors in
 CC Japan with high s-GTP levels, and were found to be different from
 CC previously reported NANB hepatitis viruses. The DNA can be used as a
 CC reagent for detecting the NANB hepatitis viral gene. The polypeptides can
 CC be used as reagents for detecting anti-NANB hepatitis antibodies or as a
 CC NANB hepatitis viral vaccine. (Updated on 25-MAR-2003 to correct PN
 CC field.) (Updated on 27-AUG-2003 to correct OS field.)
 XX SQ Sequence 547 AA;

Query Match 70.2%; Score 2144; DB 2; Length 547;
 Best Local Similarity 75.5%; Pred. No. 4.4e-197;
 Matches 397; Conservative 55; Mismatches 74; Indels 0; Gaps 0;
 QY 49 ASLRQKQVTFDRLOVLDHYRDVLKEMKAKSTVKALLSVEEACKLTPPHSAKSGFYG 108
 Db 1 ASLRQKQVTFDRLOVLDHYRDVLKEMKAKSTVKALLSVEEACKLTPPHSAKSGFYG 60
 QY 109 AKOVRSLSSRAVNHRSVWKLLEDVDTPIQTIMAKNEVFCVQPEKGRKPARLIVFPD 168
 Db 61 AKOVRSLSSRAVNHRSVWKLLEDVDTPIQTIMAKNEVFCVQPEKGRKPARLIVFPD 120
 QY 169 LGVRVCEKMALYDVVSTLQAVMGSSYGQYSPKORVEFLVNTWKAKKCPMGFSYDTRCF 228
 Db 121 LGVRVCEKMALYDVVSTLQAVMGSSYGQYSPKORVEFLVNTWKAKKCPMGFSYDTRCF 180
 QY 229 DSTVTENDIRVESIYQCCDLAPPEARQAIRSLTERLYVGGPMTNSKGQYRCRASGV 288
 Db 181 DSTVTENDIRVESIYQCCDLAPPEARQAIRSLTERLYVGGPMTNSKGQYRCRASGV 240
 QY 289 LTTSCGNTLTCYLKAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAAASLRFVTEAMT 348
 Db 241 FTSMGNTMTCYIKALAAACRAAGIKDPIMLVCGDDLVVISESQNGNEEDERNLRAFTAMT 300
 QY 349 RYSAPPGDPPQPEYDLELITSCSSNVSVADHAGSKRVYVLTTRDPTVPLARAATWETARHTP 408
 Db 301 RYSAPPGDPPQPEYDLELITSCSSNVSVADHAGSKRVYVLTTRDPTVPLARAATWETARHTP 360
 QY 409 VNSWLGNIIMYAPTLWARMLMTHFFSILLAOBLEKALDCQIYGACYSIEPLDLPOIIE 468
 Db 361 VNSWLGNIIMYAPTLWARMLMTHFFSILLAOBLEKALDCQIYGACYSIEPLDLPOIIE 420
 QY 469 RLHGLSAPSLHYSYSGEINRVASCLKGLGVPLVRVWRHARSVRAKLLSQGGRAATCGKY 528
 Db 421 RLHGLSAPSLHYSYSGEINRVASCLKGLGVPLVRVWRHARSVRAKLLSQGGRAATCGKY 480
 QY 529 LFNWAVRTKLKLTPIPAASRLDLSGWFGVAGYSGGDIYHLSRARPR 574
 Db 481 LFNWAVRTKLKLTPIPAASRLDLSGWFGVAGYSGGDIYHLSRARPR 526

RESULT 15
 AAR29906
 ID AAR29906 standard; protein; 389 AA.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2005, 14:47:55 ; Search time 43 Seconds
(without alignments)
1297.808 Million cell updates/sec

Title: US-10-712-479-4

Perfect score: 3055

Sequence: 1 MASMSYTWGALITPCAA.....GGDIYHLSLRARPRHHHH 580

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 248019

Minimum DB seq length: 0

Maximum DB seq length: 585

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_79.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1777	58.2	365	2 JQ0879	NS5 protein - hepa
2	1655	54.2	365	2 JQ0880	NS5 protein - hepa
3	951	31.1	189	2 S60587	non-structural pro
4	774	25.3	284	2 P50104	genome polyprotein
5	744	24.4	259	2 P50102	genome polyprotein
6	667	21.8	135	2 S44214	genome polyprotein
7	630	20.6	156	2 D39109	genome polyprotein
8	569	18.6	113	2 PCL278	NS5 protein - hepa
9	563	18.4	113	2 PCL277	NS5 protein - hepa
10	562	18.4	113	2 PCL274	NS5 protein - hepa
11	562	18.4	113	2 PCL275	NS5 protein - hepa
12	560	18.3	113	2 PCL276	NS5 protein - hepa
13	513	16.8	113	2 PCL279	NS5 protein - hepa
14	426	13.9	113	2 PCL282	NS5 protein - hepa
15	426	13.9	113	2 PCL275	polyprotein - hepa
16	425	13.9	113	2 PCL276	polyprotein - hepa
17	424	13.9	113	2 PCL281	NS5 protein - hepa
18	422	13.8	113	2 PCL283	NS5 protein - hepa
19	409	13.4	113	2 PCL280	NS5 protein - hepa
20	408	13.4	113	2 PCL273	polyprotein - hepa
21	406	13.3	113	2 PCL274	polyprotein - hepa
22	402	13.2	78	2 PS0085	hypothetical prote
23	336	11.0	266	2 P00393	genome polyprotein
24	300	9.8	132	2 P00396	genome polyprotein
25	298	9.8	132	2 P00394	genome polyprotein
26	296	9.7	60	2 S44215	genome polyprotein
27	264	8.6	71	2 PCL300	genome polyprotein
28	261	8.5	71	2 PCL302	genome polyprotein
29	260	8.5	71	2 PCL301	genome polyprotein

30	225	7.4	52	2 P00555	nonstructural prot
31	224	7.3	52	2 P00558	nonstructural prot
32	223	7.3	52	2 P00556	nonstructural prot
33	221	7.2	52	2 P00554	nonstructural prot
34	215	7.0	52	2 P00557	nonstructural prot
35	131	4.3	535	2 B49600	probable polymeras
36	128.5	4.2	532	1 RRVQC2	RNA-directed RNA p
37	123.5	4.0	581	2 JQ1879	hypothetical 65.4k
38	106	3.5	317	2 AC3410	peptidylprolyl iso
39	105.5	3.5	356	2 G96806	thauartin-like pro
40	103	3.4	315	2 T16359	hypothetical prote
41	96.5	3.2	546	2 B40407	sterol carrier pro
42	94	3.1	535	2 AE3202	hypothetical prote
43	92.5	3.0	565	2 T47423	hypothetical prote
44	91	3.0	523	2 JC7556	linoleoyl-CoA desa
45	90	2.9	476	2 AE2126	pyruvate kinase li

ALIGNMENTS

RESULT 1

JQ0879

NS5 protein - hepatitis C virus (strain J4) (fragment)

C:Species: hepatitis C virus

C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jul-2004

C:Accession: JQ0879

R:Okamoto, H.

submitted to JIPID, January 1991

A:Description: The 5'-terminal and 3'-terminal sequences of the genomic RNA of hepatitis

A:Reference number: JQ0879

A:Accession: JQ0879

A:Molecule type: genomic RNA

A:Residues: 1-365 <OKA>

A:Cross-references: UNIPROT:Q81717

A:Experimental source: strain J4

C:Superfamily: hepatitis C virus genome polyprotein

Query Match 58.2%; Score 1777; DB 2; Length 365;

Best Local Similarity 96.8%; Pred. No. 7e-129;

Matches 333; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

Qy	231	TVTENDIRVESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNGCYRRCRASGVLT	290
Db	1	TVTESDIRVESIYQCCDLAPEARQAIRSLTERLYVGGPLTNSKGQNGCYRRCRASGVLT	60
Qy	291	TSCGNLTCTYLKAAACRAAKLQDCTMLVNGDDLVIKESAGTOEDAAASLAVFTTEAMTRY	350
Db	61	TSCGNLTCTYLKATAACRAAKLQDCTMLVNGDDLVIKESAGTOEDAAALRVFTTEAMTRY	120
Qy	351	SAPPGDPPQPPYDLELITSCSSNVSAHDAGKRVYVLTTRDPTVPLARAANETARHTPN	410
Db	121	SAPPGDPPQPPYDLELITSCSSNVSAHDAGKRVYVLTTRDPTVPLARAANETARHTPN	180
Qy	411	SWLGNIIWYAPTLWARMILMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLPQIERL	470
Db	181	SWLGNIIWYAPALWARMILMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLPQIERL	240
Qy	471	HGLSAFSLHSSPGEINRVASCLRLKGLVPPPLRVWRHARSVRAKLLSQGGRAAICGKYL	530
Db	241	HGLSAFSLHSSPGEINRVASCLRLKGLVPPPLRVWRHARSVRAKLLSQGGRAATCGRYL	300
Qy	531	NWAVRTKLTPIPAASRLDLSGMFVAGYSGGDIYHLSLRARPR	574
Db	301	NWAVRTKLTPIPAASRLDLSGMFVAGYSGGDIYHLSLRARPR	344

RESULT 2

JQ0880

NS5 protein - hepatitis C virus (strain J1) (fragment)

C:Species: hepatitis C virus

C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jul-2004

C:Accession: JQ0880

R;Okamoto, H.
submitted to JIPID, January 1991
A:Description: The 5'-terminal and 3'-terminal sequences of the genomic RNA of hepatitis
A:Reference number: JQ0879
A:Accession: JQ0880
A:Molecule type: genomic RNA
A:Residues: 1-365 <OKA>
A:Cross-references: UNIPROT:Q81716
A:Experimental source: strain J1
C:Superfamily: hepatitis C virus genome polyprotein

Query Match 54.2%; Score 1655; DB 2; Length 365;
Best Local Similarity 88.7%; Pred. No. 1.7e-119;
Matches 305; Conservative 22; Mismatches 17; Indels 0; Gaps 0;

Qy 231 TVTENDIRVESIYQCCDLAPAPARQAIRSLTERLYVGGPMTNSGQNGYRRCRASGVLT 290
Db 1 TVTESDIRTEAIYQCCDLDPQARVAIKSLTERLYVGGPLTNSRGNGYRRCRASGVLT 60

Qy 291 TSCGNTLTCTYLKAAACRAAKLODCTMLVNGDDLWVICESAGTQEDAAASLRVFTFMTTRY 350
Db 61 TSCGNTLTCTYLKAAACRAAGLRDCTMLVCGDDLWVICESQGVQEDAAASLRVFTFMTTRY 120

Qy 351 SAPGCDPPQPEYDLELITSCSNVSVAHDAGSKRVYLTTRDPTVPLARAAWETARHTPVN 410
Db 121 SAPGCDPPQPEYDLELITSCSNVSVAHDGTRKRVYLTTRDPTVPLARAAWETARHTPVN 180

Qy 411 SWLGNIMYAPTLWARMILMTHFFSILIAQBOLEKALDCQIYGACYSIEPLDLPQIIRL 470
Db 181 SWLGNIMYAPTLWARMILMTHFFSILIAQBOLEKALDCQIYGACYSIEPLDLPQIIRL 240

Qy 471 HGLSAFSLHSYSGEINRVASCLKLGVPPLVVRHRRARSVRAKLLSQGGRAAICGKYL 530
Db 241 HGLSAFSLHSYSGEINRVAACTRLKLGVPPLVVRHRRARSVRALLSRGGRAAICGKYL 300

Qy 531 NWAVRTKLTPIPAASRLDLSGMVFAGYSGGDIYHLSRPR 574
Db 301 NWAVRTKLTPIAAAGRLDLSGHTAGYSGGDIYHNSVSHRPR 344

RESULT 3
S60587
non-structural protein NS5 - hepatitis C virus (fragment)
C:Species: hepatitis C virus
C:Date: 19-Mar-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004
C:Accession: S60587
R;Xu, L.Z.; Martinot-Peignoux, M.; Marcellin, P.; Benhamou, J.P.; Larzul, D.
J. Hepatol. 20, 598-602, 1994
A:Title: Comparison of the sensitivity of nested PCR in the 5' non-coding and the NS5 re
A:Reference number: S60587; MUID:94351175; PMID:8071535
A:Accession: S60587
A:Status: preliminary; translation not shown
A:Molecule type: genomic RNA
A:Residues: 1-189 <XUL>
A:Cross-references: UNIPROT:Q68959; EMBL:X59934; NID:g556891; PIDN:CAA42557.1; PID:g4388
C:Superfamily: hepatitis C virus genome polyprotein

Query Match 31.1%; Score 951; DB 2; Length 189;
Best Local Similarity 93.7%; Pred. No. 9.4e-66;
Matches 177; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Qy 153 PEKGRKPARLIVFDLGRVCEKALYDVVSTLPQAVMGSSYGFQVSPKQVFEFLVNTW 212
Db 1 PEKGRKPARLIVFDLGRVCEKALYDVVSTLPQAVMGSSYGFQVSPKQVFEFLVNAW 60

Qy 213 KAKKCPMGFSYDTRCFDSTVTENDIRVEESIQCCDLAPAPARQAIRSLTERLYVGGPMTN 272
Db 61 KSKKCPMGFAIDTRCFDSTVTENDIRVEESIQCCDLAPAPARQIRSLTERLYIGGLTN 120

Qy 273 SKGQNGYRRCRASGVLTSCGNTLTCTYLKAAACRAAKLODCTMLVNGDDLWVICESAG 332
Db 121 SKGQNGYRRCRASGVLTSCGNTLTCTYLKAAACRAAKLODCTMLVCGDDLWVICESGG 180

Qy 333 TOEDAASLR 341
Db 181 TOEDAASLR 189

RESULT 4
PS0104
genome polyprotein - hepatitis C virus (isolate pt) (fragments)
N:Contains: NS5 protein; structural protein
C:Species: hepatitis C virus
A:Note: host Pan troglodytes (chimpanzee)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004
C:Accession: PS0104
R;Takeuchi, K.; Boonmar, S.; Kubo, Y.; Katayama, T.; Harada, H.; Ohbayashi, A.; Choo, Q.I.
Gene 91, 287-291, 1990
A:Title: Hepatitis C viral cDNA clones isolated from a healthy carrier donor implicated i
A:Reference number: PS0102; MUID:91007289; PMID:2170237
A:Accession: PS0104
A:Molecule type: genomic RNA
A:Residues: 1-284 <TAK>
A:Cross-references: UNIPROT:Q81633; GB:D90078; GB:D90079
A:Note: the authors translated the codon TGT for residue 274 as Lys
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: polyprotein
F:1-129/Product: structural protein (fragment) #status predicted <XPR>
F:130-284/Product: NS5 protein (fragment) #status predicted <NS5>

Query Match 25.3%; Score 774; DB 2; Length 284;
Best Local Similarity 88.6%; Pred. No. 6.6e-52;
Matches 140; Conservative 13; Mismatches 5; Indels 0; Gaps 0;

Qy 378 HDASGKRVYLTTRDPTVPLARAAWETARHTPVNSWLGNIMYAPTLWARMILMTHFFSIL 437
Db 127 HKFAGKRVYLTTRDPTVPLARAAWETARHTPVNSWLGNIMYAPTLWARMILMTHFFSVL 186

Qy 438 LAOBOLEKALDCQIYGACYSIEPLDLPQIIRLHGLSAFSLHSYSGEINRVASCLKLG 497
Db 187 IARDQLEQALDCQIYGACYSIEPLDLPQIIRLHGLSAFSLHSYSGEINRVAACTRLKLG 246

Qy 498 VPPLVVRHRRARSVRAKLLSQGGRAAICGKYLFWAVR 535
Db 247 VPPLVVRHRRARSVRALLSRGGRAAICGKYLFWAVR 284

RESULT 5
PS0102
genome polyprotein - hepatitis C virus (isolate J1) (fragments)
N:Contains: NS5 protein; structural protein
C:Species: hepatitis C virus
A:Note: host Homo sapiens (man)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004
C:Accession: PS0102
R;Takeuchi, K.; Boonmar, S.; Kubo, Y.; Katayama, T.; Harada, H.; Ohbayashi, A.; Choo, Q.I.
Gene 91, 287-291, 1990
A:Title: Hepatitis C viral cDNA clones isolated from a healthy carrier donor implicated i
A:Reference number: PS0102; MUID:91007289; PMID:2170237
A:Contents: J1
A:Accession: PS0102
A:Molecule type: genomic RNA
A:Residues: 1-259 <TAK>
A:Cross-references: UNIPROT:Q81617; GB:D90077; GB:D90080
A:Note: the authors translated the codon TGT for residue 255 as Lys
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: polyprotein
F:1-117/Product: structural protein (fragment) #status predicted <SPR>
F:118-259/Product: NS5 protein (fragment) #status predicted <NS5>

Query Match 24.4%; Score 744; DB 2; Length 259;
Best Local Similarity 99.3%; Pred. No. 1.2e-49;
Matches 141; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 388 LTRDPTVPLARAAWETARHTPVNSWLGNIMYAPTLWARMILMTHFFSILIAQBOLEKAL 447

Db 118 LTRDPTVPLARAWEETARHTPNVSNLGNITMYAPTLMWMLMTHTFFSILLAQEQLEKAL 177
QY 448 DCQIYGACYSIEPLDLPOIIRLHGLSAFSLHSYSGPEINRVASCLRLKLGVPPLRVWRHR 507
Db 178 DCQIYGACYSIEPLDLPOIIRLHGLSAFSLHSYSGPEINRVASCLRLKLGVPPLRVWRHR 237
QY 508 ARSVRAKLLSQGGRAAICGKYL 529
Db 238 ARSVRAKLLSQGGRAATCGKYL 259

RESULT 6
S44214
genome polyprotein - hepatitis C virus (fragment)
N:Contains: NS5a protein
C:Species: hepatitis C virus
C>Date: 13-Jan-1995 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C:Accession: S44214
R:Feucht, H.H.
submitted to the EMBL Data Library, April 1994
A:Reference number: S44213
A:Accession: S44214
A:Molecule type: mRNA
A:Residues: 1-135 <F&U>
A:Cross-references: UNIPROT:Q68962; EMBL:X78954; NID:9475176; PIDN:CAA55551.1; PID:G9398
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: polyprotein
F:1-135/Product: NS5a protein #status predicted <MAT>

Query Match 21.8%; Score 667; DB 2; Length 135;
Best Local Similarity 91.9%; Pred. No. 4.1e-44;
Matches 124; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 147 EVFCVQPEKGGKPARLIVFPDLGVRCCKMALYDVSTLPQAVNGSSYGFQYSPQQRVE 206
Db 1 EVFCVQPEKGGKPARLIVFPDLGVRCCKMALYDVSTLPQAVNGSSYGFQYSPQQRVE 60

QY 207 FLVNTWKAICPMGFSYDTRCFDSTVTENDIRVEESIYQCCLAPAPARQAIRSLTERLYV 266
Db 61 FLVNAWKSQKPMGFAYCTRCFDSTVTESDIRVEESIYQCCLAPAPARQAIRSLTERLYV 120

QY 267 GGPMTNSKGQNGYR 281
Db 121 GGPLTNSKGQNGYR 135

RESULT 7
D39109
genome polyprotein - hepatitis C virus (fragments)
C:Species: hepatitis C virus
C>Date: 18-Oct-1991 #sequence_revision 18-Oct-1991 #text_change 17-Nov-2000
C:Accession: D39109
R:Han, J.H.; Shyanala, V.; Richman, K.H.; Brauer, M.J.; Irvine, B.; Urdea, M.S.; Tekamp-
Proc. Natl. Acad. Sci. U.S.A. 88, 1711-1715, 1991
A:Title: Characterization of the terminal regions of hepatitis C viral RNA: identification
A:Reference number: A39109; MUID:91156678; PMID:1705704
A:Accession: D39109
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-156 <HAN>
A:Cross-references: GB:M58406; GB:M58407
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: polyprotein

Query Match 20.6%; Score 630; DB 2; Length 156;
Best Local Similarity 89.2%; Pred. No. 3.5e-41;
Matches 116; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 445 KALDCQIYGACYSIEPLDLPOIIRLHGLSAFSLHSYSGPEINRVASCLRLKLGVPPLRVW 504
Db 6 KALDCQIYGACYSIEPLDLPOIIRLHGLSAFSLHSYSGPEINRVASCLRLKLGVPPLRVW 65

QY 505 RHRARSVRKLLSQGGRAAICGKYLFWNAVRTKLTPTIPAASRLDLSGWFVAGYSGGDI 564

Db 66 RHRARSVRARLLARGGAAICGKYLFWNAVRTKLTPTIPAAGOLDLSGWFAGYSGGDI 125
QY 565 YHLSRARPR 574
Db 126 YHSVSHARPR 135

RESULT 8
PC1278
NS5 protein - hepatitis C virus (strain KI-4) (fragment)
C:Species: hepatitis C virus
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: PC1278
R:Enomoto, N.; Tanaka, A.; Nakao, T.; Date, T.
Biochem. Biophys. Res. Commun. 170, 1021-1025, 1990
A:Title: There are two major types of hepatitis C virus in Japan.
A:Reference number: PC1274; MUID:90358793; PMID:2117923
A:Accession: PC1278
A:Molecule type: mRNA
A:Residues: 1-113 <ENO>
A:Cross-references: UNIPROT:Q01187; GB:D10645; GB:D90554; NID:g221668; PIDN:BAA01492.1; I
C:Genetics:
A:Gene: NS5
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein

Query Match 18.6%; Score 569; DB 2; Length 113;
Best Local Similarity 95.6%; Pred. No. 1.1e-36;
Matches 108; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 230 STVTENDIRVEESIYQCCLAPAPARQAIRSLTERLYVGGPMTNSKGQNGYRRCRASGVL 289
Db 1 STVTENDIRVEESIYQCCLAPAPARQAIRSLTERLYVGGPMTNSKGQNGYRRCRASGVL 60

QY 290 TTSCGNTLTCTYKAAACRAAKLQDCTMLVNGDDLWVICSAGTQEDAAASLRV 342
Db 61 TTSCGNTLTCTYKAAACRAAKLQDCTMLVNGDDLWVICSAGTQEDAAASLRV 113

RESULT 9
PC1277
NS5 protein - hepatitis C virus (strain KI-3) (fragment)
C:Species: hepatitis C virus
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: PC1277
R:Enomoto, N.; Tanaka, A.; Nakao, T.; Date, T.
Biochem. Biophys. Res. Commun. 170, 1021-1025, 1990
A:Title: There are two major types of hepatitis C virus in Japan.
A:Reference number: PC1274; MUID:90358793; PMID:2117923
A:Accession: PC1277
A:Molecule type: mRNA
A:Residues: 1-113 <ENO>
A:Cross-references: UNIPROT:Q070640; GB:D10644; GB:D90553; NID:g221666; PIDN:BAA01491.1; I
C:Genetics:
A:Gene: NS5
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein

Query Match 18.4%; Score 563; DB 2; Length 113;
Best Local Similarity 94.7%; Pred. No. 3.2e-36;
Matches 107; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 230 STVTENDIRVEESIYQCCLAPAPARQAIRSLTERLYVGGPMTNSKGQNGYRRCRASGVL 289
Db 1 STVTENDIRVEESIYQCCLAPAPARQAIRSLTERLYVGGPMTNSKGQNGYRRCRASGVL 60

QY 290 TTSCGNTLTCTYKAAACRAAKLQDCTMLVNGDDLWVICSAGTQEDAAASLRV 342
Db 61 TTSCGNTLTCTYKAAACRAAKLQDCTMLVNGDDLWVICSAGTQEDAAASLRV 113

RESULT 10

PC1274
NS5 protein - hepatitis C virus (strain KI) (fragment)
C:Species: hepatitis C virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: PC1274
R:Enomoto, N.; Tanaka, A.; Nakao, T.; Date, T.
Biochem. Biophys. Res. Commun. 170, 1021-1025, 1990
A:Title: There are two major types of hepatitis C virus in Japan.
A:Reference number: PC1274; MUID:90358793; PMID:2117923
A:Accession: PC1274
A:Molecule type: mRNA
A:Residues: 1-113 <ENO>
A:Cross-references: UNIPROT:Q01192; GB:D10641; GB:D90550; NID:g221658; PIDN:BAA01488.1;
C:Genetics:
A:Gene: NS5
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein

Query Match 18.4%; Score 562; DB 2; Length 113;
Best Local Similarity 95.6%; Pred. No. 3.8e-36;
Matches 108; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 230 STVTENDIRVEESIQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNGCYRRCRAGSVL 289
Db 1 STVTENDIRVEESIQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNGCYRRCRAGSVL 60

Qy 290 TTSCGNTLTCYLKAAACRAAKLQDCTMLVNGDDLVLVICSAGTQEDAASLRV 342
Db 61 TTSCGNTLTCYLKATAACRAAKLQDCTMLVCGDDLVLVICSAGTQEDAASLRV 113

RESULT 11
PC1275
NS5 protein - hepatitis C virus (strain KI-1) (fragment)
C:Species: hepatitis C virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: PC1275
R:Enomoto, N.; Tanaka, A.; Nakao, T.; Date, T.
Biochem. Biophys. Res. Commun. 170, 1021-1025, 1990
A:Title: There are two major types of hepatitis C virus in Japan.
A:Reference number: PC1274; MUID:90358793; PMID:2117923
A:Accession: PC1275
A:Molecule type: mRNA
A:Residues: 1-113 <ENO>
A:Cross-references: UNIPROT:Q01184; GB:D10642; GB:D90551; NID:g221662; PIDN:BAA01489.1;
C:Genetics:
A:Gene: NS5
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein

Query Match 18.4%; Score 562; DB 2; Length 113;
Best Local Similarity 94.7%; Pred. No. 3.8e-36;
Matches 107; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 230 STVTENDIRVEESIQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNGCYRRCRAGSVL 289
Db 1 STVTENDIRVEESIQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNGCYRRCRAGSVL 60

Qy 290 TTSCGNTLTCYLKAAACRAAKLQDCTMLVNGDDLVLVICSAGTQEDAASLRV 342
Db 61 TTSCGNTLTCYLKASACRAAKLQDCTMLVCGDDLVLVICSAGTQEDAANLRV 113

RESULT 12
PC1276
NS5 protein - hepatitis C virus (strain KI-2) (fragment)
C:Species: hepatitis C virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: PC1276
R:Enomoto, N.; Tanaka, A.; Nakao, T.; Date, T.
Biochem. Biophys. Res. Commun. 170, 1021-1025, 1990
A:Title: There are two major types of hepatitis C virus in Japan.
A:Reference number: PC1274; MUID:90358793; PMID:2117923

A:Accession: PC1276
A:Molecule type: mRNA
A:Residues: 1-113 <ENO>
A:Cross-references: UNIPROT:Q01185; GB:D10643; GB:D90552; NID:g221664; PIDN:BAA01490.1; I
C:Genetics:
A:Gene: NS5
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein

Query Match 18.3%; Score 560; DB 2; Length 113;
Best Local Similarity 93.8%; Pred. No. 5.4e-36;
Matches 106; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 230 STVTENDIRVEESIQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNGCYRRCRAGSVL 289
Db 1 STVTESDIRVEESIQCCDLAPEARQVIRSLTERLYIGGPLTNSKGQNGCYRRCRAGSVL 60

Qy 290 TTSCGNTLTCYLKAAACRAAKLQDCTMLVNGDDLVLVICSAGTQEDAASLRV 342
Db 61 TTSCGNTLTCYLKASACRAAKLQDCTMLVCGDDLVLVICSAGTQEDAASLRV 113

RESULT 13
PC1279
NS5 protein - hepatitis C virus (strain PT-1) (fragment)
C:Species: hepatitis C virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: PC1279
R:Enomoto, N.; Tanaka, A.; Nakao, T.; Date, T.
Biochem. Biophys. Res. Commun. 170, 1021-1025, 1990
A:Title: There are two major types of hepatitis C virus in Japan.
A:Reference number: PC1274; MUID:90358793; PMID:2117923
A:Accession: PC1279
A:Molecule type: mRNA
A:Residues: 1-113 <ENO>
A:Cross-references: UNIPROT:Q01193; GB:D10646; GB:D90555; NID:g221670; PIDN:BAA01493.1; I
C:Genetics:
A:Gene: NS5
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein

Query Match 16.8%; Score 513; DB 2; Length 113;
Best Local Similarity 85.7%; Pred. No. 2.2e-32;
Matches 96; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

Qy 230 STVTENDIRVEESIQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNGCYRRCRAGSVL 289
Db 1 STVTESDIRTEAIIYQCCDLDPQARVAIKSLTERLYVGGPLTNSRGNGCYRRCRAGSVL 60

Qy 290 TTSCGNTLTCYLKAAACRAAKLQDCTMLVNGDDLVLVICSAGTQEDAASLR 341
Db 61 TTSCGNTLTCYIKARAAACRAAGLRDCTMLVCGDDLVLVICSAGVQEDAASLR 112

RESULT 14
PC1282
NS5 protein - hepatitis C virus (strain K2b) (fragment)
C:Species: hepatitis C virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: PC1282
R:Enomoto, N.; Tanaka, A.; Nakao, T.; Date, T.
Biochem. Biophys. Res. Commun. 170, 1021-1025, 1990
A:Title: There are two major types of hepatitis C virus in Japan.
A:Reference number: PC1274; MUID:90358793; PMID:2117923
A:Accession: PC1282
A:Molecule type: mRNA
A:Residues: 1-113 <ENO>
A:Cross-references: UNIPROT:Q01189; GB:D10649; GB:D90558; NID:g221676; PIDN:BAA01496.1; I
C:Genetics:
A:Gene: NS5
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein

Search completed: September 22, 2005, 14:54:47
Job time : 44 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2005, 14:47:32 ; Search time 175 Seconds
(without alignments)
1697.177 Million cell updates/sec

Title: US-10-712-479-4

Perfect score: 3055

Sequence: 1 MASMSMTWTGALITPCAA.....GGDIYHSLSRAPRHHHH 580

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1438883

Minimum DB seq length: 0

Maximum DB seq length: 585

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2873	94.0	571	2 Q6B4Q1	Q6B4q1 hepatitis c
2	2870	93.9	571	2 Q66NA8	Q66na8 hepatitis c
3	2857	93.5	571	2 Q66N95	Q66n95 hepatitis c
4	2857	93.5	571	2 Q66NA6	Q66na6 hepatitis c
5	2856	93.5	571	2 Q66N84	Q66n84 hepatitis c
6	2853	93.4	571	2 Q6B4Q2	Q6B4q2 hepatitis c
7	2851	93.3	571	2 Q66N79	Q66n79 hepatitis c
8	2846	93.2	571	2 Q66N83	Q66n83 hepatitis c
9	2839	92.9	571	2 Q66NA2	Q66na2 hepatitis c
10	2837	92.9	571	2 Q66N78	Q66n78 hepatitis c
11	2837	92.9	571	2 Q66NA0	Q66na0 hepatitis c
12	2836	92.8	571	2 Q66N74	Q66n74 hepatitis c
13	2820	92.3	571	2 Q66N89	Q66n89 hepatitis c
14	2803	91.8	571	2 Q66NB0	Q66nb0 hepatitis c
15	2679	87.7	571	2 Q66NA7	Q66na7 hepatitis c
16	2673	87.5	571	2 Q66N75	Q66n75 hepatitis c
17	2673	87.5	571	2 Q66NA5	Q66na5 hepatitis c
18	2672	87.5	571	2 Q66N86	Q66n86 hepatitis c
19	2672	87.5	571	2 Q66N93	Q66n93 hepatitis c
20	2672	87.5	571	2 Q66NA4	Q66na4 hepatitis c
21	2671	87.4	571	2 Q66N99	Q66n99 hepatitis c
22	2669	87.4	571	2 Q66N91	Q66n91 hepatitis c
23	2668	87.3	571	2 Q66N83	Q66n83 hepatitis c
24	2665	87.2	571	2 Q66N94	Q66n94 hepatitis c
25	2664	87.2	571	2 Q66N98	Q66n98 hepatitis c
26	2663	87.2	571	2 Q66NA3	Q66na3 hepatitis c
27	2661	87.1	571	2 Q66N82	Q66n82 hepatitis c
28	2661	87.1	571	2 Q66NA1	Q66na1 hepatitis c
29	2661	87.1	571	2 Q66N95	Q66n95 hepatitis c
30	2660	87.1	571	2 Q66N85	Q66n85 hepatitis c
31	2660	87.1	571	2 Q66N90	Q66n90 hepatitis c

32 2660 87.1 571 2 Q66N96 hepatitis c
33 2660 87.1 571 2 Q66NA9 hepatitis c
34 2659 87.0 571 2 Q66N81 hepatitis c
35 2658 87.0 571 2 Q66N80 hepatitis c
36 2657 87.0 571 2 Q66N82 hepatitis c
37 2656 86.9 571 2 Q66N88 hepatitis c
38 2655 86.9 571 2 Q66N76 hepatitis c
39 2655 86.9 571 2 Q66N87 hepatitis c
40 2650 86.7 571 2 Q66N77 hepatitis c
41 2646 86.6 571 2 Q66N73 hepatitis c
42 2646 86.6 571 2 Q66N92 hepatitis c
43 2644 86.5 571 2 Q66NB1 hepatitis c
44 2644 86.5 571 2 Q66NB4 hepatitis c
45 2636 86.3 571 2 Q66N97 hepatitis c

ALIGNMENTS

RESULT 1
Q6B4Q1 PRELIMINARY; PRT; 571 AA.
AC Q6B4Q1;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE NS5b (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY682462; AAT84167.1; --
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003688; F:RNA-directed RNA polymerase activity; IEA.
DR GO: GO:0006350; P:transcription; IEA.
DR GO: GO:0019079; P:viral genome replication; IEA.
DR InterPro: IPR002166; HCV_RDRP.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF00998; Viral_RDRP; 1.
FT NON_TER 1
FT NON_TER 571
SQ SEQUENCE 571 AA; 63223 MW; 975788F1925D0DBC CRC64;

Query Match 94.0%; Score 2873; DB 2; Length 571;
Best Local Similarity 95.8%; Pred. No. 6.9e-218;
Matches 543; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

Qy 5 SMSYTWGALITPCAAEESKLPIINALSNLRLHNLVYSTSRASLRQKVTDFRLOVL 64
Db 5 SMSYTWGALITPCAAEESKLPIINALSNLRLHNLVYSTSRASLRQKVTDFRLOVL 64
Qy 65 DDHYRDVLKEMKAKASTVKAALLSVEACKLTPPHSAKSKFGYGAQDVRSLSRAVNHIR 124
Db 65 DDHYRDVLKEMKAKASTVKAALLSVEACKLTPPHSAKSKFGYGAQDVRSLSRAVNHIR 124
Qy 125 SVWKDLLEDTDTPITQTTIMAKNEVFCVQPEKGGKRPARLIVFPDLGVRVCEKMALYDVVS 184
Db 125 SVWKDLLEDTDTPITQTTIMAKNEVFCVQPEKGGKRPARLIVFPDLGVRVCEKMALYDVVS 184
Qy 185 TLPOAVMGSSYGFQSPQKQRFVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESYI 244
Db 185 TLPOAVMGSSYGFQSPQKQRFVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESYI 244
Qy 245 QCCLDLAPEARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTCGNTLTCLKAA 304

```
Db 245 QCCDLAPARQAIIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTSCGNTLTCYLKAT 304
Qy 305 AACRAAKLQDCITMLVNGDDLVVICESAGTQDEASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
Db 305 AACRAAKLQDCITMLVNGDDLVVICESAGTQDEASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
Qy 365 ELITSCSSNVSAHDAAGKRVYILTRDPTPLARAAMETARHTPVNSWLGNIIIMYAPTLW 424
Db 365 ELITSCSSNVSAHDAAGKRVYILTRDPTPLARAAMETARHTPVNSWLGNIIIMYAPTLW 424
Qy 425 ARMLIMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
Db 425 ARMLIMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
Qy 485 EINRVASCLRLKGLVPPPLRVWRHARSVRAKLLSQGGRAAICGKYLFWNAVTRTKLTPIP 544
Db 485 EINRVASCLRLKGLVPPPLRVWRHARSVRAKLLSQGGRAAICGKYLFWNAVTRTKLTPIP 544
Qy 545 AASRLDLGSGWFGVAGYSGGDIYHSLRA 571
Db 545 AASRLDLGSGWFGVAGYSGGDIYHSLRA 571
```

RESULT 2

```
Q66NA8 PRELIMINARY; PRT; 571 AA.
AC Q66NA8
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY682754; AAU08311.1; -.
DR InterPro: IPR002166; HCV_RDRP.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF00998; Viral_RDRP; 1.
KW Polyprotein.
KW NON_TER
FT NON_TER 1 571
SQ SEQUENCE 571 AA; 63201 MW; 5590C9F59B69B2A0 CRC64;
```

```
Query Match 93.9%; Score 2870; DB 2; Length 571;
Best Local Similarity 95.8%; Pred. No. 1.2e-217;
Matches 543; Conservative 14; Mismatches 10; Indels 0; Gaps 0;
```

```
Qy 5 SMSYTWGALITPCAABESKLPINALSNLRLHNLVYTTSRASLRQKKVTFDRQLVL 64
Db 5 SMSYTWGALITPCAABETKLPINALSNLRLHNLVYTTSRASLRQKKVTFDRQLVL 64
Qy 65 DDHYRDVLKEMKAKASTVKALLSVEEACKLTPPHSAKSKFGYCAKDVRSLSRAVNHIR 124
Db 65 DDHYRDVLKEMKAKASTVKALLSVEEACKLTPPHSAKSKFGYCAKDVRSLSRAVNHIR 124
Qy 125 SVWKDLLEDTDTPITQTTIMAKNEVFCVQPEKGGKRPARLIVFPDLGVRVCEKMALYDVVS 184
Db 125 SVWKDLLEDTDTPITQTTIMAKNEVFCVQPEKGGKRPARLIVFPDLGVRVCEKMALYDVVS 184
Qy 185 TLPOAVMGSSYGFQYSPQKRVFELVNTWKAICKPMGFSYDTRCFDSTVTENDIRVESIY 244
Db 185 TLPOAVMGSSYGFQYSPQKRVFELVNTWKAICKPMGFSYDTRCFDSTVTENDIRVESIY 244
```

```
Qy 245 QCCDLAPARQAIIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTSCGNTLTCYLKAA 304
Db 245 QCCDLAPARQAIIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTSCGNTLTCYLKAT 304
Qy 305 AACRAAKLQDCITMLVNGDDLVVICESAGTQDEASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
Db 305 AACRAAKLQDCITMLVNGDDLVVICESAGTQDEASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
Qy 365 ELITSCSSNVSAHDAAGKRVYILTRDPTPLARAAMETARHTPVNSWLGNIIIMYAPTLW 424
Db 365 ELITSCSSNVSAHDAAGKRVYILTRDPTPLARAAMETARHTPVNSWLGNIIIMYAPTLW 424
Qy 425 ARMLIMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
Db 425 ARMLIMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
Qy 485 EINRVASCLRLKGLVPPPLRVWRHARSVRAKLLSQGGRAAICGKYLFWNAVTRTKLTPIP 544
Db 485 EINRVASCLRLKGLVPPPLRVWRHARSVRAKLLSQGGRAAICGKYLFWNAVTRTKLTPIP 544
Qy 545 AASRLDLGSGWFGVAGYSGGDIYHSLRA 571
Db 545 AASRLDLGSGWFGVAGYSGGDIYHSLRA 571
```

RESULT 3

```
Q66N95 PRELIMINARY; PRT; 571 AA.
AC Q66N95
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY682768; AAU08324.1; -.
DR InterPro: IPR002166; HCV_RDRP.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF00998; Viral_RDRP; 1.
KW Polyprotein.
KW NON_TER
FT NON_TER 1 571
SQ SEQUENCE 571 AA; 63249 MW; F76482BE94526FDA CRC64;
```

```
Query Match 93.5%; Score 2857; DB 2; Length 571;
Best Local Similarity 94.9%; Pred. No. 1.3e-216;
Matches 536; Conservative 17; Mismatches 12; Indels 0; Gaps 0;
```

```
Qy 5 SMSYTWGALITPCAABESKLPINALSNLRLHNLVYTTSRASLRQKKVTFDRQLVL 64
Db 5 SMSYTWGALITPCAABESKLPINALSNLRLHNLVYTTSRASLRQKKVTFDRQLVL 64
Qy 65 DDHYRDVLKEMKAKASTVKALLSVEEACKLTPPHSAKSKFGYCAKDVRSLSRAVNHIR 124
Db 65 DDHYRDVLKEMKAKASTVKALLSVEEACKLTPPHSAKSKFGYCAKDVRSLSRAVNHIR 124
Qy 125 SVWKDLLEDTDTPITQTTIMAKNEVFCVQPEKGGKRPARLIVFPDLGVRVCEKMALYDVVS 184
Db 125 SVWKDLLEDTDTPITQTTIMAKNEVFCVQPEKGGKRPARLIVFPDLGVRVCEKMALYDVVS 184
Qy 185 TLPOAVMGSSYGFQYSPQKRVFELVNTWKAICKPMGFSYDTRCFDSTVTENDIRVESIY 244
Db 185 TLPOAVMGSSYGFQYSPQKRVFELVNTWKAICKPMGFSYDTRCFDSTVTENDIRVESIY 244
```

```

Qy 245 QCCDLAPARQAIRSLTERLYVGGPMTNSKQNCQYRCRASGVLTTS CGNTLT CYLKA 304
Db 245 QCCDLAPARQAIRSLTERLYVGGPMTNSKQNCQYRCRASGVLTTS CGNTLT CYLKA 304
Qy 305 AACRAAKLQDCTMLVNGDDLVI CESAGTQEDAAASLRVFTTEAMTRYSA PPDPPEYDL 364
Db 305 AACRAAKLQDCTMLVNGDDLVI CESAGTQEDAAASLRVFTTEAMTRYSA PPDPPEYDL 364
Qy 365 ELITSCSSNVSAHDASGRVYVLTTRDPTVPLARA AEWETARHTPVNSWLGNIIMYAPTLM 424
Db 365 ELITSCSSNVSAHDASGRVYVLTTRDPTVPLARA AEWETARHTPVNSWLGNIIMYAPTLM 424
Qy 425 ARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIERLHGLSAPFSLHSYSPG 484
Db 425 ARMVLTTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIERLHGLSAPFSLHSYSPG 484
Qy 485 EINRVASCLRLKGVPPPLRVWRHRARSVRAKLLSQGGR AAI CGKYLFWNAVRTKLTPTIP 544
Db 485 EINRVASCLRLKGVPPPLRVWRHRARSVRAKLLSQGGR AAI CGKYLFWNAVRTKLTPTIP 544
Qy 545 AASRLDLSGWFVAGYSGGDIYHSLSRA 571
Db 545 AASQLDLSGWFVAGYSGGDIYHVSVA 571

```

```

RESULT 4
Q66NA6
ID Q66NA6 PRELIMINARY; PRT; 571 AA.
AC Q66NA6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY682756; AAU08313.1; -
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR Pfam; PF00998; RNA_pol_PSVir.
DR Polyprotein.
KW NON TER 1 571
FT NON TER 571 571
SQ SEQUENCE 571 AA; 63183 MW; 445C9E3B09394146 CRC64;

Query Match 93.5%; Score 2857; DB 2; Length 571;
Best Local Similarity 95.2%; Pred. No. 1.3e-216;
Matches 540; Conservative 15; Mismatches 12; Indels 0; Gaps 0;

Qy 5 SMSYTWGALITPCAABESKLPINALSNLSLRHNHLYSTTSRSASLRQKVKVTFDRLOVL 64
Db 5 SMSYTWGALITPCAABESKLPINALSNLSLRHNHLYSTTSRSASLRQKVKVTFDRLOVL 64
Qy 65 DDHYRDVLKEMKAKASTVKALLSVEACKLTPPHSAKSKFGYKADVRSLSSRAVNHIR 124
Db 65 DDHYRDVLKEMKAKASTVKALLSVEACKLTPPHSAKSKFGYKADVRSLSSRAVNHIR 124
Qy 125 SVWKDLLEDTDTPITQTTIMAKNEVFCVQPEKGRKPARLI VFPDLGVRVCCKMALYDVVS 184
Db 125 SVWKDLLEDTDTPITQTTIMAKNEVFCVQPEKGRKPARLI VFPDLGVRVCCKMALYDVVS 184
Qy 185 TLPOAVMGSSYGFQYSPKQRFVFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 244

```

```

Db 185 TLPOAVMGSSYGFQYSPKQRFVFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
Qy 245 QCCDLAPARQAIRSLTERLYVGGPMTNSKQNCQYRCRASGVLTTS CGNTLT CYLKA 304
Db 245 QCCDLAPARQAIRSLTERLYVGGPMTNSKQNCQYRCRASGVLTTS CGNTLT CYLKA 304
Qy 305 AACRAAKLQDCTMLVNGDDLVI CESAGTQEDAAASLRVFTTEAMTRYSA PPDPPEYDL 364
Db 305 AACRAAKLQDCTMLVNGDDLVI CESAGTQEDAAASLRVFTTEAMTRYSA PPDPPEYDL 364
Qy 365 ELITSCSSNVSAHDASGRVYVLTTRDPTVPLARA AEWETARHTPVNSWLGNIIMYAPTLM 424
Db 365 ELITSCSSNVSAHDASGRVYVLTTRDPTVPLARA AEWETARHTPVNSWLGNIIMYAPTLM 424
Qy 425 ARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIERLHGLSAPFSLHSYSPG 484
Db 425 ARMVLTTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIERLHGLSAPFSLHSYSPG 484
Qy 485 EINRVASCLRLKGVPPPLRVWRHRARSVRAKLLSQGGR AAI CGKYLFWNAVRTKLTPTIP 544
Db 485 EINRVASCLRLKGVPPPLRVWRHRARSVRAKLLSQGGR AAI CGKYLFWNAVRTKLTPTIP 544
Qy 545 AASRLDLSGWFVAGYSGGDIYHSLSRA 571
Db 545 AASQLDLSGWFVAGYSGGDIYHVSVA 571

```

```

RESULT 5
Q66N84
ID Q66N84 PRELIMINARY; PRT; 571 AA.
AC Q66N84;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY682779; AAU08335.1; -
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR Pfam; PF00998; RNA_pol_PSVir.
DR Polyprotein.
KW NON TER 1 571
FT NON TER 571 571
SQ SEQUENCE 571 AA; 63393 MW; CD0FE085358BEE67 CRC64;

Query Match 93.5%; Score 2856; DB 2; Length 571;
Best Local Similarity 95.1%; Pred. No. 1.5e-216;
Matches 539; Conservative 17; Mismatches 11; Indels 0; Gaps 0;

Qy 5 SMSYTWGALITPCAABESKLPINALSNLSLRHNHLYSTTSRSASLRQKVKVTFDRLOVL 64
Db 5 SMSYTWGALITPCAABESKLPINALSNLSLRHNHLYSTTSRSASLRQKVKVTFDRLOVL 64
Qy 65 DDHYRDVLKEMKAKASTVKALLSVEACKLTPPHSAKSKFGYKADVRSLSSRAVNHIR 124
Db 65 DDHYRDVLKEMKAKASTVKALLSVEACKLTPPHSAKSKFGYKADVRSLSSRAVNHIR 124
Qy 125 SVWKDLLEDTDTPITQTTIMAKNEVFCVQPEKGRKPARLI VFPDLGVRVCCKMALYDVVS 184
Db 125 SVWKDLLEDTDTPITQTTIMAKNEVFCVQPEKGRKPARLI VFPDLGVRVCCKMALYDVVS 184
Qy 185 TLPOAVMGSSYGFQYSPKQRFVFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 244

```

Db 185 TLPQAVMGSSYGFQYSPQVRVEFLVNAWKSCKCPMGFPAYDTRCFDSTVTESDIRVEESIY 244
Qy 245 QCCDLAPPEARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTSCGNTLTCYLKAA 304
Db 245 QCCDLAPPEARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTSCGNTLTCYLKAT 304
Qy 305 AACRAAKLODCTMLVNGDDLVVICESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
Db 305 AACRAAKLODCTMLVCGDDLVVICESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
Qy 365 ELITSCSSNVSAHDASGKRVYILTRDPTVPLARAAMETARHTPVNSWLGNIIMVAPTLW 424
Db 365 ELITSCSSNVSAHDATGKRVYILTRDPTVPLARAAMETARHTPVNSWLGNIIMVAPTLW 424
Qy 425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSFAFSLHSYSPG 484
Db 425 ARMILMTHFFSILLAQEQLEKALDCQIYGACHSIEPLDLPQIIORLHGLSFAFSLHSYSPG 484
Qy 485 EINRVASCLRLKGVPPPLRVWHRARSVRAKLLSOGGAAICGKYLFWNAVRTKLTPIIP 544
Db 485 EINRVASCLRLKGVPPPLRVWHRARSVRAKLLSOGGAAICGKYLFWNAVRTKLTPIIP 544
Qy 545 AASRLDLSGMFVAGYSGGDIYHLSLSRA 571
Db 545 AASRLDLSGMFVAGYSGGDIYHLSVSHA 571

RESULT 6

Q6B4Q2 ID Q6B4Q2 PRELIMINARY; PRT; 571 AA.
AC Q6B4Q2
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE NS5b (Fragment).
OS Hepatitis C virus.
OC Viruses; asRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY682461; AAT94166.1; -
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0003723; P:RNA binding; IEA.
DR GO; GO:0003968; P:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF00998; Viral_RdRP; 1.
FT NON TER 1
FT NON TER 571
SQ SEQUENCE 571 AA; 63123 MW; 931743E7C7819F36 CRC64;

Query Match 93.4%; Score 2853; DB 2; Length 571;
Best Local Similarity 95.2%; Pred. No. 2.6e-216;
Matches 540; Conservative 15; Mismatches 12; Indels 0; Gaps 0;

Qy 5 SMSYTWITGALITPCAAEESKLPINALNSLRHNLVYTTSSRASLRQKQKVTDFDLQVL 64
Db 5 SMSYTWITGALITPCAAEESKLPINALNSLRHNLVYTTSSRASLRQKQKVTDFDLQVL 64
Qy 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTTPPHSAKSKFGYGAQDVRSLSRAVNHIR 124
Db 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTTPPHSAKSKFGYGAQDVRSLSRAVNHIR 124

Qy 125 SVMKDLLEDTDTRTQTTIMAKNEVFCVQPEKGRKPARLIIVFPDLGVRVCEKMALYDVVS 184
Db 125 SVMKDLLEDTATPDTTITIMAKNEVFCVQPEKGRKPARLIIVFPDLGVRVCEKMALYDVVS 184
Qy 185 TLPQAVMGSSYGFQYSPQVRVEFLVNAWKSCKCPMGFPAYDTRCFDSTVTESDIRVEESIY 244
Db 185 TLPQAVMGSSYGFQYSPQVRVEFLVNAWKSCKCPMGFPAYDTRCFDSTVTESDIRVEESIY 244
Qy 245 QCCDLAPPEARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTSCGNTLTCYLKAA 304
Db 245 QCCDLAPPEARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTSCGNTLTCYLKAS 304
Qy 305 AACRAAKLODCTMLVNGDDLVVICESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
Db 305 AACRAAKLODCTMLVCGDDLVVICESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
Qy 365 ELITSCSSNVSAHDASGKRVYILTRDPTVPLARAAMETARHTPVNSWLGNIIMVAPTLW 424
Db 365 ELITSCSSNVSAHDATGKRVYILTRDPTVPLARAAMETARHTPVNSWLGNIIMVAPTLW 424
Qy 425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSFAFSLHSYSPG 484
Db 425 ARMILMTHFFSILLAQEQLEKALDCQIYGACHSIEPLDLPQIIORLHGLSFAFSLHSYSPG 484
Qy 485 EINRVASCLRLKGVPPPLRVWHRARSVRAKLLSOGGAAICGKYLFWNAVRTKLTPIIP 544
Db 485 EINRVASCLRLKGVPPPLRVWHRARSVRAKLLSOGGAAICGKYLFWNAVRTKLTPIIP 544
Qy 545 AASRLDLSGMFVAGYSGGDIYHLSLSRA 571
Db 545 AASRLDLSGMFVAGYSGGDIYHLSVSHA 571

RESULT 7

Q66N79 ID Q66N79 PRELIMINARY; PRT; 571 AA.
AC Q66N79
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; asRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY682784; AAU08340.1; -
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF00998; Viral_RdRP; 1.
FT NON TER 1
FT NON TER 571
SQ SEQUENCE 571 AA; 63317 MW; 4BC8AC1E98153857 CRC64;

Query Match 93.3%; Score 2851; DB 2; Length 571;
Best Local Similarity 95.2%; Pred. No. 3.7e-216;
Matches 540; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

Qy 5 SMSYTWITGALITPCAAEESKLPINALNSLRHNLVYTTSSRASLRQKQKVTDFDLQVL 64
Db 5 SMSYTWITGALITPCAAEESKLPINALNSLRHNLVYTTSSRASLRQKQKVTDFDLQVL 64
Qy 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTTPPHSAKSKFGYGAQDVRSLSRAVNHIR 124
Db 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTTPPHSAKSKFGYGAQDVRSLSRAVNHIR 124

Qy 125 SVWKOLLEDTTPIOTTIMAKNEVFCVQPEKGRKPARLIVFPDGLGVRVCEKMAIYDVVS 184
 Db 125 SVWKOLLEDTTPIOTTIMAKNEVFCVQPEKGRKPARLIVFPDGLGVRVCEKMAIYDVVS 184
 Qy 185 TLPOAVMGSSYGFQSPQKORVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
 Db 185 TLPOAVMGSSYGFQSPQKORVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
 Qy 245 OCCDLAPARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTCGNTLTCLYKAA 304
 Db 245 OCCDLAPARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTCGNTLTCLYKAT 304
 Qy 305 AACRAAKLQDCTMLVNGDDLVIICISAGTQEDAAASLRVFTAMTRYSAAPPDPPPEYDL 364
 Db 305 AACRAAKLQDCTMLVNGDDLVIICISAGTQEDAAASLRVFTAMTRYSAAPPDPPPEYDL 364
 Qy 365 ELITSCSNVSAVHADSGKRVVYLTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTILW 424
 Db 365 ELITSCSNVSAVHADSGKRVVYLTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTILW 424
 Qy 425 ARMITLMTFFSILLAOEALDQCIYGCYSIEPLDLPOLIERLHGLSFAFSLHSYSPG 484
 Db 425 ARMITLMTFFSILLAOEALDQCIYGCYSIEPLDLPOLIERLHGLSFAFSLHSYSPG 484
 Qy 485 EINRVASCLRLKGLVPPPLRVWRHARSVRAKLLSQGGRAAICGKYLFWNAVRTKLTPIIP 544
 Db 485 EINRVASCLRLKGLVPPPLRVWRHARSVRAKLLSQGGRAAICGKYLFWNAVRTKLTPIIP 544
 Qy 545 AASRLDLSGWFVAGYSGGDIYHLSRA 571
 Db 545 DASRLDLSGWFVAGYSGGHIYHVSVA 571

RESULT 8
 Q66N83 ID Q66N83 PRELIMINARY; PRT; 571 AA.
 AC Q66N83; DT 25-OCT-2004 (TremBLrel. 28, Created)
 DT 25-OCT-2004 (TremBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TremBLrel. 28, Last annotation update)
 DE Polyprotein (fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
 RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
 RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
 RT HCV/HIV Coinfected Subjects."
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY682780; AAU08336.1; -;
 DR InterPro; IPR002166; HCV RdRP.
 DR InterPro; IPR007095; RNA pol DS PS.
 DR InterPro; IPR007094; RNA pol_PSVir.
 DR Pfam; PF00998; Viral_RdRP; 1.
 KW Polyprotein.
 FT NON TER 1
 FT NON TER 571
 SQ SEQUENCE 571 AA; 63177 MW; 8649F9B5ED28DEED CRC64;

Query Match 93.2%; Score 2846; DB 2; Length 571;
 Best Local Similarity 94.9%; Pred. No. 9.3e-216;
 Matches 538; Conservative 17; Mismatches 12; Indels 0; Gaps 0;
 Qy 5 SMSYTWTCALITPCAABESKLPINALNSLRHNLVYVSTTSRGSASLRQKVTDFRLQVL 64
 Db 5 SMSYTWTCALITPCAABESKLPINALNSLRHNLVYVSTTSRGSASLRQKVTDFRLQVL 64
 Qy 65 DDHYRDVLKEMKAKASTVKAALLSVEEACKLTPPHSAKSKFGYGAOVRSLSRAVNHIR 124

Db 65 DDHYRDVLKEMKAKASTVKAALLSVEEACKLTPPHSAKSKFGYGAOVRSLSRAVNHIR 124
 Qy 125 SVWKOLLEDTTPIOTTIMAKNEVFCVQPEKGRKPARLIVFPDGLGVRVCEKMAIYDVVS 184
 Db 125 SVWKOLLEDTTPIOTTIMAKNEVFCVQPEKGRKPARLIVFPDGLGVRVCEKMAIYDVVS 184
 Qy 185 TLPOAVMGSSYGFQSPQKORVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
 Db 185 TLPOAVMGSSYGFQSPQKORVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
 Qy 245 OCCDLAPARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTCGNTLTCLYKAA 304
 Db 245 OCCDLAPARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTCGNTLTCLYKAS 304
 Qy 305 AACRAAKLQDCTMLVNGDDLVIICISAGTQEDAAASLRVFTAMTRYSAAPPDPPPEYDL 364
 Db 305 AACRAAKLQDCTMLVNGDDLVIICISAGTQEDAAASLRVFTAMTRYSAAPPDPPPEYDL 364
 Qy 365 ELITSCSNVSAVHADSGKRVVYLTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTILW 424
 Db 365 ELITSCSNVSAVHADSGKRVVYLTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTILW 424
 Qy 425 ARMITLMTFFSILLAOEALDQCIYGCYSIEPLDLPOLIERLHGLSFAFSLHSYSPG 484
 Db 425 ARMITLMTFFSILLAOEALDQCIYGCYSIEPLDLPOLIERLHGLSFAFSLHSYSPG 484
 Qy 485 EINRVASCLRLKGLVPPPLRVWRHARSVRAKLLSQGGRAAICGKYLFWNAVRTKLTPIIP 544
 Db 485 EINRVASCLRLKGLVPPPLRVWRHARSVRAKLLSQGGRAAICGKYLFWNAVRTKLTPIIP 544
 Qy 545 AASRLDLSGWFVAGYSGGDIYHLSRA 571
 Db 545 AASQLDLSGWFVAGYSGGDIYHSGSHA 571

RESULT 9
 Q66N83 ID Q66N83 PRELIMINARY; PRT; 571 AA.
 AC Q66N83; DT 25-OCT-2004 (TremBLrel. 28, Created)
 DT 25-OCT-2004 (TremBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TremBLrel. 28, Last annotation update)
 DE Polyprotein (fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
 RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
 RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
 RT HCV/HIV Coinfected Subjects."
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY682760; AAU08317.1; -;
 DR InterPro; IPR002166; HCV RdRP.
 DR InterPro; IPR007095; RNA pol DS PS.
 DR InterPro; IPR007094; RNA pol_PSVir.
 DR Pfam; PF00998; Viral_RdRP; 1.
 KW Polyprotein.
 FT NON TER 1
 FT NON TER 571
 SQ SEQUENCE 571 AA; 63227 MW; F54BDE550ACDEEBC CRC64;

Query Match 92.9%; Score 2839; DB 2; Length 571;
 Best Local Similarity 94.5%; Pred. No. 3.3e-215;
 Matches 536; Conservative 19; Mismatches 12; Indels 0; Gaps 0;
 Qy 5 SMSYTWTCALITPCAABESKLPINALNSLRHNLVYVSTTSRGSASLRQKVTDFRLQVL 64
 Db 5 SMSYTWTCALITPCAABESKLPINALNSLRHNLVYVSTTSRGSASLRQKVTDFRLQVL 64
 Qy 65 DDHYRDVLKEMKAKASTVKAALLSVEEACKLTPPHSAKSKFGYGAOVRSLSRAVNHIR 124

```
Db 65 DDHYRDVLKEIKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAQVNRSLSSKATNHIR 124
Qy 125 SVWKDLLEDDTPTQTTIMAKNEVFCVQPEKGGKRPAPRLIVFPDLGVRVCEKMALYDVVS 184
Db 125 SVWKDLLEDDTPTQTTIMAKNEVFCVQPEKGGKRPAPRLILFPDLGVRVCEKMALYDVVS 184
Qy 185 TLPOAVNGSSYGFQYSPQGRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
Db 185 TLPOAVNGSSYGFQYSPQGRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
Qy 245 QCCDLAPARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTCGNTLTCYLKAA 304
Db 245 QCCDLAPARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTCGNTLTCYLKAS 304
Qy 305 AACRAAKLQDCTMLVNGDDLVIICESAGTQDAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
Db 305 AACRAAKLQDCTMLVNGDDLVIICESAGTQDAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
Qy 365 ELITSCSSNVSAHDASGRVYVLTTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTLW 424
Db 365 ELITSCSSNVSAHDASGRVYVLTTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTLW 424
Qy 425 ARMLTMTFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIRLHGLSFAFSLHSYSPG 484
Db 425 ARMLTMTFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIRLHGLSFAFSLHSYSPG 484
Qy 485 EINRVASCLRLGVPPLVRWHRARSVRAKLLSOGGRAAICGKYLFWNAVRTKLTPIIP 544
Db 485 EINRVASCLRLGVPPLVRWHRARSVRAKLLSOGGRAAICGKYLFWNAVRTKLTPIIP 544
Qy 545 AASRLDLSGWFVAGYSGGDIYHLSRA 571
Db 545 AASQLDLSGWFVAGYSGGDIYHVSVA 571
```

RESULT 10

```
Q66N78 ID Q66N78 PRELIMINARY; PRT; 571 AA.
AC Q66N78;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY682785; AAU08341.1; -
DR InterPro; IPR002166; HCV_RDRP.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR Pfam; PF00998; RNA_pol_PSVir.
KW Polyprotein.
FT NON_TER 1 571
FT NON_TER 571 571
SQ SEQUENCE 571 AA; 63360 MW; 4521656175D8334A CRC64;
```

```
Query Match 92.9%; Score 2837; DB 2; Length 571;
Best Local Similarity 94.7%; Pred. No. 4.8e-215;
Matches 537; Conservative 17; Mismatches 13; Indels 0; Gaps 0;
Qy 5 SMSYTWTCALITPCAABESKLPINALNSLRHNLVYSTTSRSASLRQKKVTFDRLOVL 64
Db 5 SMSYTWTCALITPCAABESKLPINALNSLRHNLVYSTTSRSASLRQKKVTFDRLOVL 64
```

```
Qy 65 DDHYRDVLKEIKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAQVNRSLSSKAVNHIR 124
Db 65 DDHYRDVLKEIKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAQVNRSLSSKAVNHIR 124
Qy 125 SVWKDLLEDDTPTQTTIMAKNEVFCVQPEKGGKRPAPRLIVFPDLGVRVCEKMALYDVVS 184
Db 125 SVWKDLLEDDTPTQTTIMAKNEVFCVQPEKGGKRPAPRLIVFPDLGVRVCEKMALYDVVS 184
Qy 185 TLPOAVNGSSYGFQYSPQGRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
Db 185 TLPOAVNGSSYGFQYSPQGRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
Qy 245 QCCDLAPARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTCGNTLTCYLKAA 304
Db 245 QCCDLAPARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTCGNTLTCYLKAS 304
Qy 305 AACRAAKLQDCTMLVNGDDLVIICESAGTQDAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
Db 305 AACRAAKLQDCTMLVNGDDLVIICESAGTQDAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
Qy 365 ELITSCSSNVSAHDASGRVYVLTTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTLW 424
Db 365 ELITSCSSNVSAHDATGKRVYVLTTRDPTVPLARAAMETARHTPVNSWLGNIIMYAPTLW 424
Qy 425 ARMLTMTFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIRLHGLSFAFSLHSYSPG 484
Db 425 ARMLTMTFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIRLHGLSFAFSLHSYSPG 484
Qy 485 EINRVASCLRLGVPPLVRWHRARSVRAKLLSOGGRAAICGKYLFWNAVRTKLTPIIP 544
Db 485 EINRVASCLRLGVPPLVRWHRARSVRAKLLSOGGRAAICGKYLFWNAVRTKLTPIIP 544
Qy 545 AASRLDLSGWFVAGYSGGDIYHLSRA 571
Db 545 AASQLDLSGWFVAGYSGRHHYHVSVA 571
```

RESULT 11

```
Q66N90 ID Q66N90 PRELIMINARY; PRT; 571 AA.
AC Q66N90;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY682782; AAU08319.1; -
DR InterPro; IPR002166; HCV_RDRP.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR Pfam; PF00998; RNA_pol_PSVir.
KW Polyprotein.
FT NON_TER 1 571
FT NON_TER 571 571
SQ SEQUENCE 571 AA; 63254 MW; 475D83341A64631B CRC64;
```

```
Query Match 92.9%; Score 2837; DB 2; Length 571;
Best Local Similarity 94.5%; Pred. No. 4.8e-215;
Matches 536; Conservative 18; Mismatches 13; Indels 0; Gaps 0;
Qy 5 SMSYTWTCALITPCAABESKLPINALNSLRHNLVYSTTSRSASLRQKKVTFDRLOVL 64
Db 5 SMSYTWTCALITPCAABESKLPINALNSLRHNLVYSTTSRSASLRQKKVTFDRLOVL 64
```


Dd	5	SMSTYWTGALITPCAABESKLPINALSNSLLRHHNMVYATTSASQORCKVTFDRLOVL	64
Qy	65	DDHYRDVLKEMKAKASTVAKALLSVEEACKLTPPHSAKSFGYGAKDVRSLSSRAVNHIR	124
Dd	65	DDHYQDVLKEMKAKASTVAKALLSIEBEACKLTPPHSAKSFGYGAKDVRNLSLKAINHIR	124
Qy	125	SVWKDLLEDTDPIQTITMAKNBVCQPBEKGGRKPARIIVFPDGLGVRCCEKMALYDVVS	184
Dd	125	SVWEDLLEDTDPIQTITMAKNBVCQPBEKGGRKPARIIVFPDGLGVRCCEKMALYDVVS	184
Qy	185	TLPQAVNGSSGYFOYSPKORVFLVNTWKAKKCPMGFSYDTRCFDSVTVENTDIRVBESIY	244
Dd	185	TLPQAVNGSAIGFYQISPGQGVFLVNAWAKSRTPMGFAIDTRCFDSVTVESDIRVBESIY	244
Qy	245	QCDDLAPARQAIRSLTERLYIYGPPMTNSKGQNCYRRCRASGLVLTSCGTLTLCYLKAA	304
Dd	245	QCDDLAPARQAIRSLTERLYIGPLTNSKGQNCYRRCRASGLVLTSCGTLTLCYLKAS	304
Qy	305	AACRAAKLOCTMLVNGDDLVIJCESAGTOEDAASI RVFTTEAMTRY SAPGDPPQPEYDL	364
Dd	305	AACRAAKLOCTMLVCGDDLVIJCESAGTOEDAANLRVFTEAMTRY SAPGDPPQPEYDL	364
Qy	365	ELITS CSSNVSAHSDAGKEVVYLTDPOTVPLARA AEWETARHTPVNSWLGNIIIMYAPT LW	424
Dd	365	ELITS CSSNVSAHDATGKR VYLTDPOTVPLARA AEWETARHTPVNSWLGNIIIMYAPT LW	424
Qy	425	ARMILMTHFFSIILAEOLEKALDCOIYGACYSEIPLDLPOIERLHGLSAFSLHSYSPG	484
Dd	425	ARMILMTHFFSIILAEOLEKALDCOIYGACYSEIPLDLPOIORLHGLSAFSLHSYSPG	484
Qy	485	EINRVASCLRKLGVPPLVRVWRHARSVRKLLSQGGRAAICGKYLFNFWAVRTKLKLTPIP	544
Dd	485	EINRVASCLRKLGVPPLVRVWRHARSVRKLLSQGGRAATCGRYLFNFWAVRTKLKLT PIS	544
Qy	545	AASRLDLSGMFVAGYSGGDIYHLSRA	571
Dd	545	AASOLDLSGMFVAGYSGGOIYHVSVA	571

```

RESULT 13
Q66N89 PRELIMINARY; PRT; 571 AA.
AC Q66N89
DT 23-OCT-2004 (TEMBLrel. 28, Created)
DT 25-OCT-2004 (TEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TEMBLrel. 28, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
[1]
RN SEQUENCE FROM N.A.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV682774; AAU08330.1; -.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF00998; Viral_RdRP; 1.
DR Polyprotein.
FT NON_TER 1. 1
FT NON_TER 571 571
SQ SEQUENCE 571 AA; 63492 MW; 111E17000D53E88C CRC64;

Query Match 92.3%; Score 2820; DB 2; Length 571;
Best Local Similarity 94.0%; Pred. No. 1e-213;
Matches 533; Conservative 20; Mismatches 14; Indels 0; Gaps 0;

QY 5 SMSYTTGALITPCAABESKLPIINALNSLLRHHLNLYVYTSRSASLRQKKVTFDRLQVL 64

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2005, 14:48:15 ; Search time 43 Seconds
(without alignments)
1006.894 Million cell updates/sec

Title: US-10-712-479-4

Perfect score: 3055

Sequence: 1 MASMSYTTGALITPCAA.....GGDIYHLSLRPRHHHHH 580

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 489056

Minimum DB seq length: 0

Maximum DB seq length: 585

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2879	94.2	576	US-09-541-990A-1	Sequence 1, Appl
2	1124	36.8	496	US-08-867-611-10	Sequence 10, Appl
3	1124	36.8	496	US-09-690-359-10	Sequence 10, Appl
4	1124	36.8	496	PCT-US92-06965A-15	Sequence 15, Appl
5	908	29.7	504	US-07-853-985A-6	Sequence 6, Appl
6	908	29.7	504	US-07-681-703B-6	Sequence 6, Appl
7	908	29.7	504	US-08-184-236-6	Sequence 6, Appl
8	908	29.7	504	US-08-407-410B-6	Sequence 6, Appl
9	908	29.7	504	US-08-485-500-6	Sequence 6, Appl
10	908	29.7	504	PCT-US91-02370-6	Sequence 6, Appl
11	908	29.7	504	PCT-US94-04174-6	Sequence 6, Appl
12	884	28.9	187	US-07-853-985A-2	Sequence 2, Appl
13	884	28.9	187	US-07-681-703B-2	Sequence 2, Appl
14	884	28.9	187	US-08-184-236-2	Sequence 2, Appl
15	884	28.9	187	US-08-407-410B-2	Sequence 2, Appl
16	884	28.9	187	US-08-485-500-2	Sequence 2, Appl
17	884	28.9	187	PCT-US91-02370-2	Sequence 2, Appl
18	884	28.9	187	PCT-US94-04174-2	Sequence 2, Appl
19	759	24.8	152	US-08-444-818-62	Sequence 62, Appl
20	748	24.5	171	US-08-444-818-58	Sequence 58, Appl
21	686	22.5	387	US-08-867-611-12	Sequence 12, Appl
22	686	22.5	387	US-09-690-359-12	Sequence 12, Appl
23	686	22.5	387	PCT-US92-06965A-17	Sequence 17, Appl
24	678	22.2	156	PCT-US94-04174-19	Sequence 19, Appl
25	621	20.3	133	US-08-444-818-60	Sequence 60, Appl
26	613.5	20.1	393	US-08-867-611-14	Sequence 14, Appl
27	613.5	20.1	393	US-09-690-359-14	Sequence 14, Appl

28	613.5	20.1	393	5	PCT-US92-06965A-19	Sequence 19, Appl
29	584	19.1	516	1	US-08-638-911A-41	Sequence 41, Appl
30	570	18.7	113	2	US-08-256-568B-82	Sequence 82, Appl
31	570	18.7	113	3	US-09-038-369B-82	Sequence 82, Appl
32	570	18.7	113	4	US-09-378-900A-82	Sequence 82, Appl
33	570	18.7	113	4	US-09-899-044-82	Sequence 82, Appl
34	570	18.7	113	4	US-09-878-281A-214	Sequence 214, Appl
35	549	18.0	113	3	US-08-836-075A-58	Sequence 58, Appl
36	536	17.5	113	3	US-08-836-075A-54	Sequence 54, Appl
37	535	17.5	113	3	US-08-836-075A-56	Sequence 56, Appl
38	531	17.4	113	3	US-08-836-075A-62	Sequence 62, Appl
39	521	17.1	113	3	US-08-836-075A-68	Sequence 68, Appl
40	515	16.9	267	3	US-08-454-928-10	Sequence 10, Appl
41	499	16.3	113	3	US-08-836-075A-64	Sequence 64, Appl
42	467	15.3	100	4	US-08-469-260A-396	Sequence 396, Appl
43	467	15.3	100	4	US-08-488-446-396	Sequence 396, Appl
44	467	15.3	100	4	US-08-467-344A-396	Sequence 396, Appl
45	467	15.3	100	4	US-08-424-550B-396	Sequence 396, Appl

ALIGNMENTS

RESULT 1
US-09-541-990A-1
; Sequence 1, Application US/09541990A
; Patent No. 6434489
; GENERAL INFORMATION:
; APPLICANT: Lesburg, Charles A.
; APPLICANT: Cable, Michael
; APPLICANT: Hong, Zhi
; APPLICANT: Mannarino, Anthony
; APPLICANT: Weber, Patricia
; TITLE OF INVENTION: Compositions of Hepatitis C Virus NS5B Polymerase and
; TITLE OF INVENTION: Methods for Crystallizing Same
; FILE REFERENCE: IN0988 US
; CURRENT APPLICATION NUMBER: US/09/541,990A
; CURRENT FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 576
; TYPE: PRT
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (335)
; OTHER INFORMATION: amino acids at 335, 344 and 550 may be threonine
; OTHER INFORMATION: or valine at 335, valine or alanine at 344 and
; OTHER INFORMATION: arginine or glutamine at 550
US-09-541-990A-1

Query Match	94.2%;	Score 2879;	DB 4;	Length 576;
Best Local Similarity	96.0%;	Pred. No. 1.2e-286;		
Matches 545;	Conservative 11;	Mismatches 12;	Indels 0;	Gaps 0;
QY	7	SYTTWGTGALITPCAAEESKLPINALNSLLRHHNLVYTTSSASLRQKKVTFDRLOVLDD	66	
Db	9	SYTTWGTGALITPCAAEESKLPINALNSLLRHHNMVYATTSSAGLRQKKVTFDRLOVLDD	68	
QY	67	HYRDVLKEMKAKASTVKAKLLSVEEACKLTTPHSAKSKFGYGAQKDVRLSSRAVNHRSV	126	
Db	69	HYRDVLKEMKAKASTVKAKLLSVEEACKLTTPHSAKSKFGYGAQKDVRLNLSKRAVNHRSV	128	
QY	127	WKDLLEDVTDTPITQTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRCCKMALYDVVYSTL	186	
Db	129	WKDLLEDVTDTPITQTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRCCKMALYDVVYSTL	188	
QY	187	PQAVMGSSYGFQYSPKQRYVEFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIYQC	246	
Db	189	PQVVMGSSYGFQYSPKQRYVEFLVNTWKSXKNPMGFSDYDTRCFDSTVTENDIRVEESIYQC	248	
QY	247	CDLAPEARQAIRSLTERLYVVGCPMTNSKGQNCGYRRRCRASGLVTTSCGNTLTLCVLKAAA	306	

Db 249 CDLAPARQAISLTERLYIGPLTNSKGNGCYRRCRASGLTTCGNTLTCTYLKASAA 308
Qy 307 CRAAKLODCTMLVNGDDLVIICESAGTOEDDAASLRVFTTEAMTRYSAAPPDPPQPEYDLEL 366
Db 309 CRAAKLODCTMLVNGDDLVIICESAGTOEDDAASLRVFTTEAMTRYSAAPPDPPQPEYDLEL 368
Qy 367 ITSCSSNVSVAHADSGKRVYLLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLPWAR 426
Db 369 ITSCSSNVSVAHADSGKRVYLLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLPWAR 428
Qy 427 MILMTHFFSILLAQOEKALDCQIYACYSIEPLDLPQIIERLHGLSAFSLHSYSGEI 486
Db 429 MILMTHFFSILLAQOEKALDCQIYACYSIEPLDLPQIIERLHGLSAFSLHSYSGEI 488
Qy 487 NRVASCLRLKLVPPRLVRVHRARSVRKLLSQGGRAAICGKYLFWNAVTKLTPIPAA 546
Db 489 NRVASCLRLKLVPPRLVRVHRARSVRKLLSQGGRAATCGKYLFWNAVTKLTPIPAA 548
Qy 547 SRLDLSCWFMVAGYSGDDIYHLSRARPR 574
Db 549 SXLDLSGWMFMVAGYSGDDIYHLSRARPR 576

RESULT 2
US-08-671-611-10
; Sequence 10, Application US/08867611
; Patent No. 6172189
; GENERAL INFORMATION:
; APPLICANT: DEVARE, SUSHIL G
; APPLICANT: CASEY, SURESH M
; APPLICANT: CASEY, JAMES M
; APPLICANT: DAILEY, STEPHEN H
; APPLICANT: DAWSON, GEORGE J
; APPLICANT: GUTIERREZ, ROBIN A
; APPLICANT: LESNIEWSKI, RICHARD R
; APPLICANT: STEWART, JAMES L
; APPLICANT: RUPPRECHT, KEVIN R
; TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES
; STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/867,611
; FILING DATE: 02-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,757
; FILING DATE:
; APPLICATION NUMBER: US/08/179,896
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/572,822
; FILING DATE: 24-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/614,069
; FILING DATE: 07-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,561
; FILING DATE: 21-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,565

; FILING DATE: 21-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,566
; FILING DATE: 21-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 4834. US.P6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-937-9556
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 496 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-867-611-10
Query Match 36.8%; Score 1124; DB 3; Length 496;
Best Local Similarity 84.8%; Pred. No. 2.4e-106;
Matches 212; Conservative 14; Mismatches 24; Indels 0; Gaps 0;
Qy 64 LDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAQDVRSLSRAVNH 123
Db 247 LDHYQDVLKVKAAASKVKANLLSVEEACSLTPPHSAKSKFGYGAQDVRSLSRAVNH 306
Qy 124 RSVWKOLLEDDTPIQTTIMAKNEVFCVQPEKGRKPARLIVFPDPLGVRVCEKALYDVV 183
Db 307 NSVWKOLLEDDTPIQTTIMAKNEVFCVQPEKGRKPARLIVFPDPLGVRVCEKALYDVV 366
Qy 184 STLPAQVNGSSYGFQYSPKORVRELNVNTWAKKCPMGFSYDTRCFDSTVTENDIRVEESI 243
Db 367 TKLPLAVNGSSYGFQYSPKORVRELNVNTWAKKCPMGFSYDTRCFDSTVTENDIRVEESI 426
Qy 244 YQCCDLAPARQAIRSLTERLYVGGPMNTNSKGNGCYRRCRASGLTTCGNTLTCTYLKA 303
Db 427 YQCCDLAPARQAIRSLTERLYVGGPMNTNSKGNGCYRRCRASGLTTCGNTLTCTYLKA 486
Qy 304 AACRAAKLQ 313
Db 487 RAACRAAGLQ 496
RESULT 3
US-09-690-359-10
; Sequence 10, Application US/09690359
; Patent No. 6593083
; GENERAL INFORMATION:
; APPLICANT: DEVARE, SUSHIL G
; APPLICANT: DESAI, SURESH M
; APPLICANT: CASEY, JAMES M
; APPLICANT: DAILEY, STEPHEN H
; APPLICANT: DAWSON, GEORGE J
; APPLICANT: GUTIERREZ, ROBIN A
; APPLICANT: LESNIEWSKI, RICHARD R
; APPLICANT: STEWART, JAMES L
; APPLICANT: RUPPRECHT, KEVIN R
; TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES
; STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25

RESULT 4
PCT-US92-06965A-15
; Sequence 15, Application PC/TUS9206965A
; GENERAL INFORMATION:
; APPLICANT: DEVARE, S.
; APPLICANT: DESAI, S.
; APPLICANT: DAILEY, S.
; TITLE OF INVENTION: HCV SYNTHETIC PEPTIDE FROM NS1 REGION
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADRESSEE: ABBOTT LABORATORIES
;

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/853.985A
; FILING DATE: 19920320
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/505,611
; FILING DATE: 06-APR-1990
; APPLICATION NUMBER: US 07/594,854
; FILING DATE: 09-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Dehlinger, Peter J.
; REGISTRATION NUMBER: 28,006
; REFERENCE/DOCKET NUMBER: 4600-0076.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-324-0880
; TELEFAX: 415-324-0960
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-853-985A-6

Query Match      29.7%; Score 908; DB 1; Length 504;
Best Local Similarity 87.2%; Pred. No. 3.9e-84;
Matches 177; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

QY 5 SMSYTWGALITPCAAEESKLPINALSNLLRHHNLVYSTTSRSASLRQKVTDFRLQVL 64
Db 302 SMSYWTGALVTPCAEEQKLPINALSNLLRHHNLVYSTTSRSACQKQKVTDFRLQVL 361
QY 65 DDHYRDVLKEMKAKASTVKALLSVEEACKLTPPHSAKSKFGYGAQDVRSLSRAVNHIR 124
Db 362 DSHYQDVLKEVKAASKVKANLLSVEEACSLTPPHSAKSKFGYGAQDVRSRAVNHIR 421
QY 125 SVWKDLLEDTPQTITIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 184
Db 422 SVWKDLLEDNVTPTDITIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVT 481
QY 185 TLPOAVMGSSYGFQYSPQKQVEF 207
Db 482 KLPLAVMGSSYGFQYSPQKQVEF 504

RESULT 6
US-07-681-703B-6
; Sequence 6, Application US/07681703B
; Patent No. 5443965
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory
; APPLICANT: Kim, Jungshuh P.
; APPLICANT: Moeckli, Randolph
; TITLE OF INVENTION: Hepatitis C Virus Epitopes
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/681,703B
; FILING DATE: 05-APR-1991

```

```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 505,611
; FILING DATE: 06-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 594,854
; FILING DATE: 09-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-076.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-681-703B-6

Query Match      29.7%; Score 908; DB 1; Length 504;
Best Local Similarity 87.2%; Pred. No. 3.9e-84;
Matches 177; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

QY 5 SMSYTWGALITPCAAEESKLPINALSNLLRHHNLVYSTTSRSASLRQKVTDFRLQVL 64
Db 302 SMSYWTGALVTPCAEEQKLPINALSNLLRHHNLVYSTTSRSACQKQKVTDFRLQVL 361
QY 65 DDHYRDVLKEMKAKASTVKALLSVEEACKLTPPHSAKSKFGYGAQDVRSLSRAVNHIR 124
Db 362 DSHYQDVLKEVKAASKVKANLLSVEEACSLTPPHSAKSKFGYGAQDVRSRAVNHIR 421
QY 125 SVWKDLLEDTPQTITIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 184
Db 422 SVWKDLLEDNVTPTDITIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVT 481
QY 185 TLPOAVMGSSYGFQYSPQKQVEF 207
Db 482 KLPLAVMGSSYGFQYSPQKQVEF 504

RESULT 7
US-08-184-236-6
; Sequence 6, Application US/08184236
; Patent No. 5538865
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory
; APPLICANT: Kim, Jungshuh P.
; APPLICANT: Moeckli, Randolph
; TITLE OF INVENTION: Hepatitis C Virus Epitopes
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: P.O. BOX 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/184,236
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/505,611
; FILING DATE: 06-APR-1990
; APPLICATION NUMBER: US 07/594,854
; FILING DATE: 09-OCT-1990

```


APPLICATION NUMBER: US 07/853,985
FILING DATE: 20-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0113
TELEPHONE: 415-324-0880
TELEFAX: 415-324-0960
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 504 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-184-236-6

Query Match 29.7%; Score 908; DB 1; Length 504;
Best Local Similarity 87.2%; Pred. No. 3.9e-84;
Matches 177; Conservative 7; Mismatches 19; Indels 0; Gaps 0;
QY 5 SMSYTWTCALITPCAABESKLPINALNSLLRHNLVSTTSRSASLRQKKVTFDRLOVL 64
DB 302 SMSYTWTCALITPCAABESKLPINALNSLLRHNLVSTTSRSACQKKVTFDRLOVL 361
QY 65 DDHYRDVLKEMKAKASTVKALLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
DB 362 DSHYQDVLKEVKAASKVKANLLSVEEACSLTPPHSAKSKFGYGAKDVRCHARKAVTHIN 421
QY 125 SVWKDLLEDTPITOTTIMAKNEVFCVQPEKGRKPARLIIVFPDLGVRVCEKMALYDVVVS 184
DB 422 SVWKDLLEDNVTPIIDTTIMAKNEVFCVQPEKGRKPARLIIVFPDLGVRVCEKMALYDVVT 481
QY 185 TLPOAVMGSSYGFQYSPKQVVEF 207
DB 482 KLPLAVMGSSYGFQYSPQQRVEF 504

RESULT 8
US-08-407-410B-6
Sequence 6, Application US/08407410B
Patent No. 5843636
GENERAL INFORMATION:
APPLICANT: Reyes, Gregory
APPLICANT: Kim, Jungshuh P.
APPLICANT: Moeckli, Randolph
APPLICANT: Simonsen, Christian C.
TITLE OF INVENTION: Hepatitis C Virus Epitopes
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Peter J. Dehlinger
STREET: 350 Cambridge Ave., Suite 100
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/407,410B
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 505,611
FILING DATE: 06-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 594,854
FILING DATE: 09-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.

REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-076.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 323-8302
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 504 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-407-410B-6

Query Match 29.7%; Score 908; DB 2; Length 504;
Best Local Similarity 87.2%; Pred. No. 3.9e-84;
Matches 177; Conservative 7; Mismatches 19; Indels 0; Gaps 0;
QY 5 SMSYTWTCALITPCAABESKLPINALNSLLRHNLVSTTSRSASLRQKKVTFDRLOVL 64
DB 302 SMSYTWTCALITPCAABESKLPINALNSLLRHNLVSTTSRSACQKKVTFDRLOVL 361
QY 65 DDHYRDVLKEMKAKASTVKALLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
DB 362 DSHYQDVLKEVKAASKVKANLLSVEEACSLTPPHSAKSKFGYGAKDVRCHARKAVTHIN 421
QY 125 SVWKDLLEDTPITOTTIMAKNEVFCVQPEKGRKPARLIIVFPDLGVRVCEKMALYDVVVS 184
DB 422 SVWKDLLEDNVTPIIDTTIMAKNEVFCVQPEKGRKPARLIIVFPDLGVRVCEKMALYDVVT 481
QY 185 TLPOAVMGSSYGFQYSPKQVVEF 207
DB 482 KLPLAVMGSSYGFQYSPQQRVEF 504

RESULT 9
US-08-485-500-6
Sequence 6, Application US/08485500
Patent No. 5843639
GENERAL INFORMATION:
APPLICANT: Reyes, Gregory
APPLICANT: Kim, Jungshuh P.
APPLICANT: Moeckli, Randolph
APPLICANT: Simonsen, Christian C.
TITLE OF INVENTION: Hepatitis C Virus Epitopes
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Peter J. Dehlinger
STREET: 350 Cambridge Ave., Suite 100
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,500
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/407,410
FILING DATE:
APPLICATION NUMBER: US 505,611
FILING DATE: 06-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 594,854
FILING DATE: 09-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-076.21
TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: (415) 323-8302
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-485-500-6

Query Match 29.7%; Score 908; DB 2; Length 504;
Best Local Similarity 87.2%; Pred. No. 3.9e-84;
Matches 177; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

QY 5 SMSYTWTCALITPCAARESKLPINALSNLLRHHNLVYTTSSASLRQKKVTFDRLOVL 64
Db 302 SMSYSWTGALVTPCAABEQKLPINALSNLLRHHNLVYTTSSACQKQKVTFRLOVL 361
QY 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAQVRSLSRAVNHIR 124
Db 362 DSHYQDVLKEVKAASKVKANLLSVEEACSLTPPHSAKSKFGYGAQVRSRAVTHIN 421
QY 125 SVWKDLLEDTPITOTTIMAKNEVFCVQPEKGRKPARLIIVFPDLGVRVCEKMALYDVWS 184
Db 422 SVWKDLLEDNVTPIIDTTIMAKNEVFCVQPEKGRKPARLIIVFPDLGVRVCEKMALYDVVT 481
QY 185 TLPOAVMGSSYGFQYSPKQVVEF 207
Db 482 KLPLAVMGSSYGFQYSPGQVVEF 504

RESULT 10
PCT-US91-02370-6
; Sequence 6, Application PC/TUS9102370
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory
; APPLICANT: Kim, Jungsu P.
; APPLICANT: Moeckli, Randolph
; APPLICANT: Simonsen, Christian C.
; TITLE OF INVENTION: Hepatitis C Virus Epitopes
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Peter J. Dehlinger
; STREET: P.O. BOX 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/02370
; FILING DATE: 19910405
; CLASSIFICATION: 435.5
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 505,611
; FILING DATE: 06-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 594,854
; FILING DATE: 09-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-076.41
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 323-8302
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear

; MOLECULE TYPE: protein
PCT-US94-04174-6

Query Match 29.7%; Score 908; DB 5; Length 504;
Best Local Similarity 87.2%; Pred. No. 3.9e-84;
Matches 177; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

QY 5 SMSYTWTCALITPCAARESKLPINALSNLLRHHNLVYTTSSASLRQKKVTFDRLOVL 64
Db 302 SMSYSWTGALVTPCAABEQKLPINALSNLLRHHNLVYTTSSACQKQKVTFRLOVL 361
QY 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAQVRSLSRAVNHIR 124
Db 362 DSHYQDVLKEVKAASKVKANLLSVEEACSLTPPHSAKSKFGYGAQVRSRAVTHIN 421
QY 125 SVWKDLLEDTPITOTTIMAKNEVFCVQPEKGRKPARLIIVFPDLGVRVCEKMALYDVWS 184
Db 422 SVWKDLLEDNVTPIIDTTIMAKNEVFCVQPEKGRKPARLIIVFPDLGVRVCEKMALYDVVT 481
QY 185 TLPOAVMGSSYGFQYSPKQVVEF 207
Db 482 KLPLAVMGSSYGFQYSPGQVVEF 504

RESULT 11
PCT-US94-04174-6
; Sequence 6, Application PC/TUS9404174
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory
; APPLICANT: Kim, Jungsu P.
; APPLICANT: Moeckli, Randolph
; TITLE OF INVENTION: HCV Immunodiagnostic Antigens and Antibodies
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Peter J. Dehlinger
; STREET: 350 Cambridge Avenue, Suite 300
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04174
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 505,611
; FILING DATE: 06-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-324-0880
; TELEFAX: 415-324-0960
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-04174-6

Query Match 29.7%; Score 908; DB 5; Length 504;
Best Local Similarity 87.2%; Pred. No. 3.9e-84;
Matches 177; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

QY 5 SMSYTWTCALITPCAARESKLPINALSNLLRHHNLVYTTSSASLRQKKVTFDRLOVL 64
Db 302 SMSYSWTGALVTPCAABEQKLPINALSNLLRHHNLVYTTSSACQKQKVTFRLOVL 361
QY 65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAQVRSLSRAVNHIR 124
Db 362 DSHYQDVLKEVKAASKVKANLLSVEEACSLTPPHSAKSKFGYGAQVRSRAVTHIN 421
QY 125 SVWKDLLEDTPITOTTIMAKNEVFCVQPEKGRKPARLIIVFPDLGVRVCEKMALYDVWS 184
Db 422 SVWKDLLEDNVTPIIDTTIMAKNEVFCVQPEKGRKPARLIIVFPDLGVRVCEKMALYDVVT 481
QY 185 TLPOAVMGSSYGFQYSPKQVVEF 207
Db 482 KLPLAVMGSSYGFQYSPGQVVEF 504
```

Db 302 SMSYWTGALVTPCAAEQKLPINALNSLLRHNLVYVTSRSACORQKVTFRLOVL 361
QY 65 DPHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
Db 362 DSHYQDVLKEVKAASKVKANLLSVEEACSLTPPHSAKSKFGYGAKDVRCHARKAVTHIN 421
QY 125 SVWKDLLEDTPPIOTTIMAKNEVFCVQPEKGRKPARLIVPPDLGVRVCEKMAIYDVVS 184
Db 422 SVWKDLLEDNVPIDTTIMAKNEVFCVQPEKGRKPARLIVPPDLGVRVCEKMAIYDVVT 481
QY 185 TLPOAVMGSSYGFQYSPKORVEF 207
Db 482 KLPLAVMGSSYGFQYSPCORVEF 504

RESULT 12
US-07-853-985A-2
; Sequence 2, Application US/07853985A
; Patent No. 5436318
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory
; APPLICANT: Kim, Jungshuh P.
; APPLICANT: Moeckli, Randolph
; TITLE OF INVENTION: Hepatitis C Virus Epitopes
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Peter J. Dehlinger
; STREET: P.O. BOX 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94306

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/853.985A
; FILING DATE: 19920320
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/505,611
; FILING DATE: 06-APR-1990
; APPLICATION NUMBER: US 07/594,854
; FILING DATE: 09-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Dehlinger, Peter J.
; REGISTRATION NUMBER: 28,006
; REFERENCE/DOCKET NUMBER: 4600-0076.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-324-0880
; TELEFAX: 415-324-0960

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 187 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-853-985A-2

Query Match 28.9%; Score 884; DB 1; Length 187;
Best Local Similarity 87.6%; Pred. No. 2.2e-82;
Matches 163; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 206 EFLVNTWKAACKPMGFSDYTRCFDSTVTENDIRVEESIYQCCDLAPEARQAIRSLTERLY 265
Db 1 EFLVQWKSCKTTPMGFSYDTRCFDSTVTESDIRTEAIIYQCCDLDPQARVAIKSLTERLY 60
QY 266 VGGPMTNSKGCNCGYRRCRASGLVTTSCGNTLTCTYLKAAAACRAAKLODCTMLVNGDDL 325
Db 61 VGGPLTNSRGNCYRRCRASGLVTTSCGNTLTCTYIKARAACRAAGLODCTMLVCGDDL 120

QY 326 VICBSAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDLELITSCSSNNVSAHDAGSKRV 385
Db 121 VICBSAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDLELITSCSSNNVSAHDAGSKRV 180
QY 386 YYLTRD 391
Db 181 YYLTR 186

RESULT 13
US-07-681-703B-2
; Sequence 2, Application US/07681703B
; Patent No. 5443965
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory
; APPLICANT: Kim, Jungshuh P.
; APPLICANT: Moeckli, Randolph
; TITLE OF INVENTION: Hepatitis C Virus Epitopes
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/681,703B
; FILING DATE: 05-APR-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 505,611
; FILING DATE: 06-APR-1990
; APPLICATION NUMBER: US 594,854
; FILING DATE: 09-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-076.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 187 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-681-703B-2

Query Match 28.9%; Score 884; DB 1; Length 187;
Best Local Similarity 87.6%; Pred. No. 2.2e-82;
Matches 163; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 206 EFLVNTWKAACKPMGFSDYTRCFDSTVTENDIRVEESIYQCCDLAPEARQAIRSLTERLY 265
Db 1 EFLVQWKSCKTTPMGFSYDTRCFDSTVTESDIRTEAIIYQCCDLDPQARVAIKSLTERLY 60
QY 266 VGGPMTNSKGCNCGYRRCRASGLVTTSCGNTLTCTYLKAAAACRAAKLODCTMLVNGDDL 325
Db 61 VGGPLTNSRGNCYRRCRASGLVTTSCGNTLTCTYIKARAACRAAGLODCTMLVCGDDL 120
QY 326 VICBSAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDLELITSCSSNNVSAHDAGSKRV 385
Db 121 VICBSAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDLELITSCSSNNVSAHDAGSKRV 180
QY 386 YYLTRD 391
Db 181 YYLTR 186

Db 181 YYLTR 186

RESULT 14

US-08-184-236-2
; Sequence 2, Application US/08184236
; Patent No. 5538865
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory
; APPLICANT: Kim, Jungsuh P.
; APPLICANT: Moeckli, Randolph
; TITLE OF INVENTION: Hepatitis C Virus Epitopes
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: P.O. BOX 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/184,236
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/505,611
; FILING DATE: 06-APR-1990
; APPLICATION NUMBER: US 07/594,854
; FILING DATE: 09-OCT-1990
; APPLICATION NUMBER: US 07/853,985
; FILING DATE: 20-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0113
; TELEPHONE: 415-324-0880
; TELEFAX: 415-324-0960
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 187 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-184-236-2

Query Match 28.9%; Score 884; DB 1; Length 187;
Best Local Similarity 87.6%; Pred. No. 2.2e-82;
Matches 163; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

Qy	206	EFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIYQCCDLAPEARQAIRSLTERLY	265
Db	1	EFLVQWKSXKTPMGFSYDTRCFDSTVTESDIRTEAIYQCCDLDPQARVAIKSLTERLY	60
Qy	266	VGGPMTNSKQNGCYRRCRASGVLTTSCGNTLTCTYLKAAACRAAKLQDCTMLVNGDDL	325
Db	61	VGGPLTNSRGENCGYRRCRASGVLTTSCGNTLTCTYIKARAACRAAGLQDCTMLVCGDDL	120
Qy	326	VICESAGTQEDAAASLRVTEAMTRYSPAPGPPQPEYDLELITSCSSNVVAHDAGKRV	385
Db	121	VICESAGVQEDAAASLRVTEAMTRYSPAPGPPQPEYDLELITSCSSNVVAHDAGKRV	180
Qy	386	YYLTRD 391	
Db	181	YYLTR 186	

RESULT 15

US-08-407-410B-2
; Sequence 2, Application US/08407410B
; Patent No. 5843636
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory
; APPLICANT: Kim, Jungsuh P.
; APPLICANT: Moeckli, Randolph
; APPLICANT: Simonsen, Christian C.
; TITLE OF INVENTION: Hepatitis C Virus Epitopes
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Peter J. Dehlinger
; STREET: 350 Cambridge Ave., Suite 100
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/407,410B
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 505,611
; FILING DATE: 06-APR-1990
; APPLICATION NUMBER: US 594,854
; FILING DATE: 09-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-076.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 323-8302
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 187 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

US-08-407-410B-2

Query Match 28.9%; Score 884; DB 2; Length 187;
Best Local Similarity 87.6%; Pred. No. 2.2e-82;
Matches 163; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

Qy	206	EFLVNTWKAKCPMGFSYDTRCFDSTVTENDIRVEESIYQCCDLAPEARQAIRSLTERLY	265
Db	1	EFLVQWKSXKTPMGFSYDTRCFDSTVTESDIRTEAIYQCCDLDPQARVAIKSLTERLY	60
Qy	266	VGGPMTNSKQNGCYRRCRASGVLTTSCGNTLTCTYLKAAACRAAKLQDCTMLVNGDDL	325
Db	61	VGGPLTNSRGENCGYRRCRASGVLTTSCGNTLTCTYIKARAACRAAGLQDCTMLVCGDDL	120
Qy	326	VICESAGTQEDAAASLRVTEAMTRYSPAPGPPQPEYDLELITSCSSNVVAHDAGKRV	385
Db	121	VICESAGVQEDAAASLRVTEAMTRYSPAPGPPQPEYDLELITSCSSNVVAHDAGKRV	180
Qy	386	YYLTRD 391	
Db	181	YYLTR 186	

Search completed: September 22, 2005, 14:55:37
Job time : 44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2005, 14:49:36 ; Search time 173 Seconds
(without alignments)
1364.549 Million cell updates/sec

Title: US-10-712-479-4

Perfect score: 3055

Sequence: 1 MASMSYTWGALITPCAA.....GGDIYHLSLRAPRRHHH 580

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1826521 seqs, 407012169 residues

Total number of hits satisfying chosen parameters: 1677200

Minimum DB seq length: 0

Maximum DB seq length: 585

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3055	100.0	580	16 US-10-712-479-4	Sequence 4, Appli
2	2958	96.8	576	17 US-10-842-046-1	Sequence 1, Appli
3	2936	96.1	578	14 US-10-211-455-1	Sequence 1, Appli
4	2936	96.1	578	16 US-10-471-164-3	Sequence 3, Appli
5	2879	94.2	576	14 US-10-170-131-1	Sequence 1, Appli
6	2734.5	89.5	545	16 US-10-471-164-4	Sequence 4, Appli
7	1391	45.5	300	9 US-09-921-397-107	Sequence 107, App
8	1283	42.0	352	17 US-10-927-520-17	Sequence 17, Appl
9	1232	40.3	260	9 US-09-921-397-109	Sequence 109, App
10	975.5	31.9	576	14 US-10-211-455-3	Sequence 3, Appli
11	915	30.0	199	9 US-09-921-397-108	Sequence 108, App

12	623	20.4	127	9	US-09-921-397-110	Sequence 110, App
13	585	19.1	137	9	US-09-921-397-106	Sequence 106, App
14	570	18.7	113	9	US-09-294-121A-82	Sequence 82, Appl
15	570	18.7	113	9	US-09-899-082A-82	Sequence 82, Appl
16	570	18.7	113	9	US-09-899-302-82	Sequence 82, Appl
17	570	18.7	113	10	US-09-899-046-214	Sequence 214, App
18	570	18.7	113	10	US-09-878-281-214	Sequence 214, App
19	570	18.7	113	10	US-09-899-044-82	Sequence 82, Appl
20	570	18.7	113	10	US-09-873-224-214	Sequence 214, App
21	570	18.7	113	16	US-10-822-711-82	Sequence 82, Appl
22	549	18.0	113	9	US-09-851-138-58	Sequence 58, Appl
23	548	17.9	120	14	US-10-292-129-17	Sequence 17, Appl
24	536	17.5	113	9	US-09-851-138-54	Sequence 54, Appl
25	535	17.5	113	9	US-09-851-138-56	Sequence 56, Appl
26	531	17.4	113	9	US-09-851-138-62	Sequence 62, Appl
27	521	17.1	113	9	US-09-851-138-68	Sequence 68, Appl
28	499	16.3	113	9	US-09-851-138-64	Sequence 64, Appl
29	492	16.1	102	9	US-09-921-397-31	Sequence 31, Appl
30	467	15.3	100	8	US-08-424-550B-396	Sequence 396, App
31	464	15.2	113	9	US-09-294-121A-92	Sequence 92, Appl
32	464	15.2	113	9	US-09-899-082A-92	Sequence 92, Appl
33	464	15.2	113	9	US-09-899-302-92	Sequence 92, Appl
34	464	15.2	113	10	US-09-899-046-160	Sequence 160, App
35	464	15.2	113	10	US-09-878-281-160	Sequence 160, App
36	464	15.2	113	10	US-09-899-044-92	Sequence 92, Appl
37	464	15.2	113	10	US-09-873-224-160	Sequence 160, App
38	464	15.2	113	16	US-10-822-711-92	Sequence 92, Appl
39	445	14.6	113	9	US-09-294-121A-89	Sequence 89, Appl
40	445	14.6	113	9	US-09-899-082A-89	Sequence 89, Appl
41	445	14.6	113	9	US-09-899-302-89	Sequence 89, Appl
42	445	14.6	113	10	US-09-899-046-113	Sequence 113, App
43	445	14.6	113	10	US-09-878-281-113	Sequence 113, App
44	445	14.6	113	10	US-09-899-044-89	Sequence 89, Appl
45	445	14.6	113	10	US-09-873-224-113	Sequence 113, App

ALIGNMENTS

RESULT 1

US-10-712-479-4
; Sequence 4, Application US/10712479
; Publication No. US20040209283A1
; GENERAL INFORMATION:
; APPLICANT: Yagi, Yoshihiko
; APPLICANT: Sheets, Michael P.
; APPLICANT: Wells, Peter A.
; APPLICANT: Shelly, John A.
; APPLICANT: Poorman, Roger A.
; APPLICANT: Epps, Dennis E.
; TITLE OF INVENTION: A Continuous-Read Assay for the Detection of De Novo HCV RNA Polyn
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 02-333-A
; CURRENT APPLICATION NUMBER: US/10/712,479
; CURRENT FILING DATE: 2003-11-13
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 580
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: C-terminally truncated HCV NS5B polymerase (C delta 21 NS5B)
US-10-712-479-4

Query Match 100.0%; Score 3055; DB 16; Length 580;
Best Local Similarity 100.0%; Pred. No. 4.1e-269;
Matches 580; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASMSYTWGALITPCAAEESKLPINALSNSLRHNLVYSTRSASLRQKVTDR 60

DB 1 MASMSYTWGALITPCAAEESKLPINALSNSLRHNLVYSTRSASLRQKVTDR 60

```
QY 61 LQVLDHVRDVLKEMKAKASTVKAKLSVEEACKLTPPHSAKSKFGYGAQDVRSLSRAV 120
DB 61 LQVLDHVRDVLKEMKAKASTVKAKLSVEEACKLTPPHSAKSKFGYGAQDVRSLSRAV 120
QY 121 NHRSVWKDLLEDTDTPQTITMAKNEVFCVQPEKGGKPKARLIVFPDLGVVRVCEKMAIY 180
DB 121 NHRSVWKDLLEDTDTPQTITMAKNEVFCVQPEKGGKPKARLIVFPDLGVVRVCEKMAIY 180
QY 181 DWSTLPOAVNGSSYGFQYSPKQRFVFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVE 240
DB 181 DWSTLPOAVNGSSYGFQYSPKQRFVFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVE 240
QY 241 ESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGONCYRRCRASGVLTTCGNTLTCTY 300
DB 241 ESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGONCYRRCRASGVLTTCGNTLTCTY 300
QY 301 LKAAACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQ 360
DB 301 LKAAACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQ 360
QY 361 EYDLELITSCSSNSVVAHDASGKRVYILTRDPTVPLARAANETARHTPVNSWLGNIIMYA 420
DB 361 EYDLELITSCSSNSVVAHDASGKRVYILTRDPTVPLARAANETARHTPVNSWLGNIIMYA 420
QY 421 PTLWARMILMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLPOIERLHGLSAFSLHS 480
DB 421 PTLWARMILMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLPOIERLHGLSAFSLHS 480
QY 481 YSPGEINRVASCLRLKGLVPPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFWNAVRTKLT 540
DB 481 YSPGEINRVASCLRLKGLVPPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFWNAVRTKLT 540
QY 541 TPIPAASRLDLSGWFVAGYSGGDIYHLSLRARPRHHHHH 580
DB 541 TPIPAASRLDLSGWFVAGYSGGDIYHLSLRARPRHHHHH 580

RESULT 2
US-10-842-046-1
; Sequence 1, Application US/10842046
; Publication No. US2005000348A1
; GENERAL INFORMATION:
; APPLICANT: Boehringer Ingelheim International GmbH
; TITLE OF INVENTION: Hepatitis C Virus NS5B Polymerase Inhibitor Binding Pocket
; FILE REFERENCE: 13/123
; CURRENT APPLICATION NUMBER: US/10/842,046
; CURRENT FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/469,604
; PRIOR FILING DATE: 2003-05-09
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 1
; LENGTH: 576
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-842-046-1

Query Match 96.8%; Score 2958; DB 17; Length 576;
Best Local Similarity 96.0%; Pred. No. 2.8e-260;
Matches 553; Conservative 16; Mismatches 7; Indels 0; Gaps 0;

QY 5 SMSYTWTCALITPCAAEESKLIPINALSNLSLRHNLVYVSTTSRSASLRQKKVTFDRLOVL 64
DB 1 SMSYTWTCALITPCAAEESKLIPINALSNLSLRHNLVYVSTTSRSASLRQKKVTFDRLOVL 60
QY 65 DDHYRDVLKEMKAKASTVKAKLSVEEACKLTPPHSAKSKFGYGAQDVRSLSRAVNHIR 124
DB 61 DDHYRDVLKEMKAKASTVKAKLSVEEACKLTPPHSAKSKFGYGAQDVRSLSRAVNHIR 120
QY 125 SVWKDLLEDTDTPQTITMAKNEVFCVQPEKGGKPKARLIVFPDLGVVRVCEKMAIYDVS 184
DB 121 SVWKDLLEDTDTPQTITMAKNEVFCVQPEKGGKPKARLIVFPDLGVVRVCEKMAIYDVS 180

QY 61 LQVLDHVRDVLKEMKAKASTVKAKLSVEEACKLTPPHSAKSKFGYGAQDVRSLSRAV 120
DB 61 LQVLDHVRDVLKEMKAKASTVKAKLSVEEACKLTPPHSAKSKFGYGAQDVRSLSRAV 120
QY 121 NHRSVWKDLLEDTDTPQTITMAKNEVFCVQPEKGGKPKARLIVFPDLGVVRVCEKMAIY 180
DB 121 NHRSVWKDLLEDTDTPQTITMAKNEVFCVQPEKGGKPKARLIVFPDLGVVRVCEKMAIY 180
QY 181 DWSTLPOAVNGSSYGFQYSPKQRFVFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVE 240
DB 181 DWSTLPOAVNGSSYGFQYSPKQRFVFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVE 240
QY 241 ESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGONCYRRCRASGVLTTCGNTLTCTY 300
DB 241 ESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGONCYRRCRASGVLTTCGNTLTCTY 300
QY 301 LKAAACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQ 360
DB 301 LKAAACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQ 360
QY 361 EYDLELITSCSSNSVVAHDASGKRVYILTRDPTVPLARAANETARHTPVNSWLGNIIMYA 420
DB 361 EYDLELITSCSSNSVVAHDASGKRVYILTRDPTVPLARAANETARHTPVNSWLGNIIMYA 420
QY 421 PTLWARMILMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLPOIERLHGLSAFSLHS 480
DB 421 PTLWARMILMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLPOIERLHGLSAFSLHS 480
QY 481 YSPGEINRVASCLRLKGLVPPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFWNAVRTKLT 540
DB 481 YSPGEINRVASCLRLKGLVPPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFWNAVRTKLT 540
QY 541 TPIPAASRLDLSGWFVAGYSGGDIYHLSLRARPRHHHHH 580
DB 541 TPIPAASRLDLSGWFVAGYSGGDIYHLSLRARPRHHHHH 580
```

```
QY 185 TLPOAVNGSSYGFQYSPKQRFVFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVESSIY 244
DB 181 TLPOAVNGSSYGFQYSPKQRFVFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVESSIY 240
QY 245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGONCYRRCRASGVLTTCGNTLTCTYLKAA 304
DB 241 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGONCYRRCRASGVLTTCGNTLTCTYLKAT 300
QY 305 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 364
DB 301 AACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAASLRVFTTEAMTRYSAAPPDPPQPEYDL 360
QY 365 ELITSCSSNSVVAHDASGKRVYILTRDPTVPLARAANETARHTPVNSWLGNIIMYAPTLM 424
DB 361 ELITSCSSNSVVAHDASGKRVYILTRDPTVPLARAANETARHTPVNSWLGNIIMYAPTLM 420
QY 425 ARMITLMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLPOIERLHGLSAFSLHSYSPG 484
DB 421 ARMITLMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLPOIERLHGLSAFSLHSYSPG 480
QY 485 EINRVASCLRLKGLVPPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFWNAVRTKLTPI 544
DB 481 EINRVASCLRLKGLVPPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFWNAVRTKLTPI 540
QY 545 AASRLDLSGWFVAGYSGGDIYHLSLRARPRHHHHH 580
DB 541 AASRLDLSGWFVAGYSGGDIYHLSLRARPRHHHHH 576

RESULT 3
US-10-211-455-1
; Sequence 1, Application US/10211455
; Publication No. US20030108862A1
; GENERAL INFORMATION:
; APPLICANT: Boehringer Ingelheim (Canada) Ltd.
; TITLE OF INVENTION: DIRECT BINDING ASSAY FOR IDENTIFYING
; TITLE OF INVENTION: INHIBITORS OF HCV POLYMERASE
; FILE REFERENCE: 13/088
; CURRENT APPLICATION NUMBER: US/10/211,455
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/310,272
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 578
; TYPE: PRT
; ORGANISM: HCV
US-10-211-455-1

Query Match 96.1%; Score 2936; DB 14; Length 578;
Best Local Similarity 95.5%; Pred. No. 2.9e-258;
Matches 552; Conservative 18; Mismatches 6; Indels 2; Gaps 1;

QY 5 SMSYTWTCALITPCAAEESKLIPINALSNLSLRHNLVYVSTTSRSASLRQKKVTFDRLOVL 64
DB 1 SMSYTWTCALITPCAAEESKLIPINALSNLSLRHNLVYVSTTSRSASLRQKKVTFDRLOVL 60
QY 65 DDHYRDVLKEMKAKASTVKAKLSVEEACKLTPPHSAKSKFGYGAQDVRSLSRAVNHIR 124
DB 61 DDHYRDVLKEMKAKASTVKAKLSVEEACKLTPPHSAKSKFGYGAQDVRSLSRAVNHIR 120
QY 125 SVWKDLLEDTDTPQTITMAKNEVFCVQPEKGGKPKARLIVFPDLGVVRVCEKMAIYDVS 184
DB 121 SVWKDLLEDTDTPQTITMAKNEVFCVQPEKGGKPKARLIVFPDLGVVRVCEKMAIYDVS 180
QY 185 TLPOAVNGSSYGFQYSPKQRFVFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVESSIY 244
DB 181 TLPOAVNGSSYGFQYSPKQRFVFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVESSIY 240
QY 245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGONCYRRCRASGVLTTCGNTLTCTYLKAA 304
DB 241 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGONCYRRCRASGVLTTCGNTLTCTYLKAS 300
```

```
Qy 305 AACRAAKLQDCTMLVNGDDLVI CESAGTQBDAAASLRVFTTEAMTRYSGAPGDPQPEYDL 364
Db 301 AACRAAKLQDCTMLVNGDDLVI CESAGTQBDAAANLRVFTTEAMTRYSGAPGDLPOPEYDL 360
Qy 365 ELITSCSSNVSAHDASGKRYVYLTTRDPTVPLARAAMETARHTPNWNLGNIIMYAPTLW 424
Db 361 ELITSCSSNVSAHDASGKRYVYLTTRDPTVPLARAAMETARHTPNWNLGNIIMYAPTLW 420
Qy 425 ARMTLMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIRLHGLSFAFSLHSYSPG 484
Db 421 ARMTLMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIRLHGLSFAFSLHSYSPG 480
Qy 485 EINRVASCLRKLGVPPLRVWRHRSVRAKLLSOGGAAICGKYLFWNAVTKLKLTPIP 544
Db 481 EINRVASCLRKLGVPPLRVWRHRSVRAKLLSOGGAAICGKYLFWNAVTKLKLTPIP 540
Qy 545 AASRLDLSGWFVAGYSGGDIYHLSRARPR--HHHHH 580
Db 541 AASRLDLSGWFVAGYSGGDIYHLSRARPRLEHHHHH 578

RESULT 4
US-10-471-164-3
; Sequence 3, Application US/10471164
; Publication No. US20040110126A1
; GENERAL INFORMATION:
; APPLICANT: George KUKOLJ and Ginette MCKERCHER
; TITLE OF INVENTION: HCV POLYMERASE INHIBITOR ASSAY
; FILE REFERENCE: 13/094
; CURRENT APPLICATION NUMBER: US/10/471.164
; CURRENT FILING DATE: 2003-09-05
; PRIOR APPLICATION NUMBER: 60/274,374
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 578
; TYPE: PRP
; ORGANISM: HCV
US-10-471-164-3

Query Match 96.1%; Score 2936; DB 16; Length 578;
Best Local Similarity 95.5%; Pred. No. 2.9e-253;
Matches 552; Conservative 18; Mismatches 6; Indels 2; Gaps 1;

Qy 5 SMSYTWGALITPCAABESKLPINALNSLLRHNLVYTSRSASLRQKVTDFRLQVL 64
Db 1 SMSYTWGALITPCAABESKLPINALNSLRHNMVYTSRSAAALRQKVTDFRLQVL 60
Qy 65 DDHYRDVLKEMKAKASTVAKLLSVEACKLTPPHSAKSKFGYGAQDVRSLSRAVNHIR 124
Db 61 DDHYRDVLKEMKAKASTVAKLLSVEACKLTPPHSAKSKFGYGAQDVRSLSRAVNHIR 120
Qy 125 SVWKDLLEDTPITOTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 184
Db 121 SVWKDLLEDTPITOTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVS 180
Qy 185 TLPOAVNGSSYGFQYSPKORVEFLVNTWAKKCPMGFSYDTRCFDSTVTENDIRVESIY 244
Db 181 TLPOAVNGSSYGFQYSPKORVEFLVNTWAKKCPMGFSYDTRCFDSTVTENDIRVESIY 240
Qy 245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTCGNTLTCYLKAA 304
Db 241 QCCDLAPEARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTCGNTLTCYLKAS 300
Qy 305 AACRAAKLQDCTMLVNGDDLVI CESAGTQBDAAASLRVFTTEAMTRYSGAPGDPQPEYDL 364
Db 301 AACRAAKLQDCTMLVNGDDLVI CESAGTQBDAAANLRVFTTEAMTRYSGAPGDLPOPEYDL 360
Qy 365 ELITSCSSNVSAHDASGKRYVYLTTRDPTVPLARAAMETARHTPNWNLGNIIMYAPTLW 424
Db 361 ELITSCSSNVSAHDASGKRYVYLTTRDPTVPLARAAMETARHTPNWNLGNIIMYAPTLW 420
```

```
Qy 425 ARMTLMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIRLHGLSFAFSLHSYSPG 484
Db 421 ARMTLMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIRLHGLSFAFSLHSYSPG 480
Qy 485 EINRVASCLRKLGVPPLRVWRHRSVRAKLLSOGGAAICGKYLFWNAVTKLKLTPIP 544
Db 481 EINRVASCLRKLGVPPLRVWRHRSVRAKLLSOGGAAICGKYLFWNAVTKLKLTPIP 540
Qy 545 AASRLDLSGWFVAGYSGGDIYHLSRARPR--HHHHH 580
Db 541 AASRLDLSGWFVAGYSGGDIYHLSRARPRLEHHHHH 578

RESULT 5
US-10-170-131-1
; Sequence 1, Application US/10170131
; Publication No. US20030171874A1
; GENERAL INFORMATION:
; APPLICANT: Lesburg, Charles A.
; APPLICANT: Cable, Michael
; APPLICANT: Hong, Zhi
; APPLICANT: Mannarino, Anthony
; APPLICANT: Weber, Patricia
; TITLE OF INVENTION: Compositions of Hepatitis C Virus NS5B Polymerase and
; FILE REFERENCE: Methods for Crystallizing Same
; FILE REFERENCE: IN0988 US
; CURRENT APPLICATION NUMBER: US/10/170.131
; CURRENT FILING DATE: 2002-06-12
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 576
; TYPE: PRP
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (335)
; OTHER INFORMATION: amino acids at 335, 344 and 550 may be threonine
; OTHER INFORMATION: or valine at 335, valine or alanine at 344 and
; OTHER INFORMATION: arginine or glutamine at 550
US-10-170-131-1

Query Match 94.2%; Score 2879; DB 14; Length 576;
Best Local Similarity 96.0%; Pred. No. 4.5e-253;
Matches 545; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

Qy 7 SYTWGALITPCAABESKLPINALNSLLRHNLVYTSRSASLRQKVTDFRLQVLDD 66
Db 9 SYTWGALITPCAABESKLPINALNSLLRHNMVYATTSRAGLRQKVTDFRLQVLDD 68
Qy 67 HYRDVLKEMKAKASTVAKLLSVEACKLTPPHSAKSKFGYGAQDVRSLSRAVNHIRSV 126
Db 69 HYRDVLKEMKAKASTVAKLLSVEACKLTPPHSAKSKFGYGAQDVRSLSRAVNHIRSV 128
Qy 127 WKOLLEDTPITOTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVSTL 186
Db 129 WKOLLEDTPITOTTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVSTL 188
Qy 187 POAVNGSSYGFQYSPKORVEFLVNTWAKKCPMGFSYDTRCFDSTVTENDIRVESIYQC 246
Db 189 POAVNGSSYGFQYSPKORVEFLVNTWAKKCPMGFSYDTRCFDSTVTENDIRVESIYQC 248
Qy 247 CDLAPEARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTCGNTLTCYLKAAA 306
Db 249 CDLAPEARQAIRSLTERLYVGGPMTNSKQNGCYRRCRASGVLTTCGNTLTCYLKAAA 308
Qy 307 CRAAKLQDCTMLVNGDDLVI CESAGTQBDAAASLRVFTTEAMTRYSGAPGDPQPEYDL 366
Db 309 CRAAKLQDCTMLVNGDDLVI CESAGTQBDAAASLRVFTTEAMTRYSGAPGDPQPEYDL 368
Qy 367 ITS CSSNVSAHDASGKRYVYLTTRDPTVPLARAAMETARHTPNWNLGNIIMYAPTLW 426
```


Db 369 ITSCSSNVSAHDASGKRVYLLTRDPTTTLARAAWETARHTPVNSWLNGLNIIMYATLWAR 428
Qy 427 MILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPOIIRLHGLSAFSLHSYSGEI 486
Db 429 MILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPOIIRLHGLSAFSLHSYSGEI 488
Qy 487 NRVASCLRLKGLVPLRVWRHRSVRKLLSQGGRAAICGKYLEFNWAVRTKLTPIPA 546
Db 489 NRVASCLRLKGLVPLRVWRHRSVRKLLSQGGRAAICGKYLEFNWAVRTKLTPIPA 548
Qy 547 SRLDLSGFWAGYSGGDIYHLSLRARPR 574
Db 549 SXLDLSGFWAGYSGGDIYHLSLRARPR 576

RESULT 6
US-10-471-164-4
; Sequence 4, Application US/10471164
; Patent No. US20040110126A1
; GENERAL INFORMATION:
; APPLICANT: George KUKOLJ and Ginette MCKERCHER
; TITLE OF INVENTION: HCV POLYMERASE INHIBITOR ASSAY
; FILE REFERENCE: 13/094
; CURRENT APPLICATION NUMBER: US/10/471.164
; CURRENT FILING DATE: 2003-09-05
; PRIOR APPLICATION NUMBER: 60/274,374
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 545
; TYPE: PRT
; ORGANISM: HCV
US-10-471-164-4

Query Match 89.5%; Score 2734.5; DB 16; Length 545;
Best Local Similarity 90.1%; Pred. No. 6.2e-240;
Matches 519; Conservative 19; Mismatches 7; Indels 31; Gaps 2;

Qy 5 SMSYTWGALITPCAAEESQIPINALSNLSLRHNLVYITTSRSASIRQKVTFDRLQVL 64
Db 1 SMSYTWGALITPCAAEESQIPINALSNLSLRHNLVYITTSRSASIRQKVTFDRLQVL 60
Qy 65 DDHYRDVLKEMKASTVAKLLSVEACKLTTPHSAKSKFGYKADVRSLSSRAVNHIR 124
Db 61 DDHYRDVLKEMKASTVAKLLSVEACKLTTPHSAKSKFGYKADVRSLSSRAVNHIR 120
Qy 125 SVWKDLLEDTDTPITQTTIMAKNEVFCVQPEKGRKPARLIIVFPDLGVRVCEKMAKYDVS 184
Db 121 SVWKDLLEDTDTPITQTTIMAKNEVFCVQPEKGRKPARLIIVFPDLGVRVCEKMAKYDVS 180
Qy 185 TLPOAVMGSSYGFQYSPKQRFVFLVNTWAKKCPMGFSYDTRCFDSTVTENDIRVESIY 244
Db 181 TLPOAVMGSSYGFQYSPKQRFVFLVNTWAKKCPMGFSYDTRCFDSTVTENDIRVESIY 240
Qy 245 QCCDLAPARQAISLTERLVVGGPMTNSKQNGYRCRCRAGVLTTCGNTLTCYLKAA 304
Db 241 QCCDLAPARQAISLTERLVVGGPMTNSKQNGYRCRCRAGVLTTCGNTLTCYLKAS 300
Qy 305 AACRAAKLQDCTMLVNGDDLVIICESAGTQDAASLRVFTTEAMTRYSGAPGDPQPEYDL 364
Db 301 AACRAAKLQDCTMLVNGDDLVIICESAGTQDAASLRVFTTEAMTRYSGAPGDLQPEYDL 360
Qy 365 ELITSCSSNVSAHDASGKRVYLLTRDPTTTLARAAWETARHTPVNSWLNGLNIIMYATLW 424
Db 361 ELITSCSSNVSAHDASGKRVYLLTRDPTTTLARAAWETARHTPVNSWLNGLNIIMYATLW 420
Qy 425 ARMLMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPOIIRLHGLSAFSLHSYSPG 484
Db 421 ARMLMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPOIIRLHGLSAFSLHSYSPG 480
Qy 485 EINRVASCLRLKGLVPLRVWRHRSVRKLLSQGGRAAICGKYLEFNWAVRTKLTPI 544

Db 481 EINRVASCLRLKGLVPLRVWRHRSVRKLLSQGGRAATCGKYLEFNWAVRTKL----- 534
Qy 545 AASRLDLSGFWAGYSGGDIYHLSLRARPRHHHHH 580
Db 535 -AAALE-----HHHHH 545

RESULT 7
US-09-921-397-107
; Sequence 107, Application US/09921397
; Patent No. US20020151484A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; TITLE OF INVENTION: SID nucleic acids and polypeptides selected from a
; TITLE OF INVENTION: pathogenic strain of the hepatitis C virus and
; TITLE OF INVENTION: applications thereof
; FILE REFERENCE: B4809A - JAZ
; CURRENT APPLICATION NUMBER: US/09/921.397
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: EP 00402225.7
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 107
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-921-397-107

Query Match 45.5%; Score 1391; DB 9; Length 300;
Best Local Similarity 87.7%; Pred. No. 7.4e-118;
Matches 263; Conservative 13; Mismatches 24; Indels 0; Gaps 0;

Qy 84 AKLLSVEEACKLTTPHSAKSKFGYKADVRSLSSRAVNHRSVWKLLDSTDPIOTTIM 143
Db 1 ANLLSVEEACSLTPPHSAKSKFGYKADVRCHARKAVAHINSVWKLLDSTDPIOTTIM 60
Qy 144 AKNEVFCVQPEKGRKPARLIIVFPDLGVRVCEKMAKYDVSITLPOAVMGSSYGFQYSPKQ 203
Db 61 AKNEVFCVQPEKGRKPARLIIVFPDLGVRVCEKMAKYDVSITLPOAVMGSSYGFQYSPKQ 120
Qy 204 RVFLVNTWAKKCPMGFSYDTRCFDSTVTENDIRVESIYQCCDLAPARQAISLTER 263
Db 121 RVFLVQAWKSKTTPMGFSYDTRCFDSTVTESDIRTEAIVQCCDLAPARQAISLTER 180
Qy 264 LYVGGPMTNSKQNGYRCRCRAGVLTTCGNTLTCYLKAAACRAAKLQDCTMLVNGDD 323
Db 181 LYVGGPLTNSRGNGCYRRRCRAGVLTTCGNTLTCYIKARAAACRAAGLQDCTMLVCGDD 240
Qy 324 LVVTCESAGTQDAASLRVFTTEAMTRYSGAPGDPQPEYDLITSCSSNVSAHDASGK 383
Db 241 LVVTCESAGVQEDAASLRVFTTEAMTRYSGAPGDPQPEYDLITSCSSNVSAHDAGK 300

RESULT 8
US-10-927-520-17
; Sequence 17, Application US/10927520
; Publication No. US20050069870A1
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New HCV clade and prototype sequences thereof
; FILE REFERENCE: 157
; CURRENT APPLICATION NUMBER: US/10/927.520
; CURRENT FILING DATE: 2004-08-27
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 352
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-10-927-520-17

Query Match 42.0%; Score 1283; DB 17; Length 352;

Query Match	40.3%	Score 1232;	DB 9;	Length 260;
Best Local Similarity	89.2%;	Pred. No. 1.9e-103;		
Matches 23;	Conservative 11;	Mismatches 17;	Indels 0;	Gaps 0;
Qy	130	LL E D T D T P I O T T I M A K N E V F C V Q B E K G K R K P A R L I V P D L G V R C E K M A L Y D V V I T L P Q A	189	
Db		:		
1	LL E D S V T P I D T T I M A K N E V F C V Q B E K G K R K P A R L I V P D L G V R C E K M A L Y D V V I S K L P L A	60		
Qy	190	V M G S S Y G F Q Y S P K Q R V E F L A V N T K A K C P M G F S Y D T R C F D S T V T V E N D I R V E S I Y Q C C D L	249	
Db	61	V M G S S Y G F Q Y S P Q R V E F L V Q A M S K K T P M G F S Y D T R C F D S T V T V E S D I R T E A I Y Q C C D L	120	
Qy	250	A P E A R Q A I R S I L T E R L Y Y G G P M T N S K G N C G Y R R C R A S G L V T T S C G N T L T C Y L K A A A A C R A	309	
Db	121	D P Q A R V A I K S I L T E R L Y Y G G P L T N S R G E N C G Y R R C R A S G L V T T S C G N T L T C Y I K A R A A C R A	180	
Qy	310	A K L Q D C T M L V N G D D L V V I C E S A G T Q E D A A S I R V F T E A M T R Y S A P P G P P Q P E Y D L E I T S	369	
Db	181	A G L Q D C T M L V C G D D L V I C E S A G V Q E D A A S I R A F T E A M T R Y S A P P G D P P Q E Y D L E I T S	240	
Qy	370	C S S N V S V A H D A S G K R V Y I L T	389	
Db	241	C S S N V S V A H D G A G K R V Y I L T	260	

233	Db	VETDIYSAKLSQDRAGIHTIARQYAGGPMIAYDGREIGYRCRSGVYTTSSSSNSLT	292
299	Qy	CYLKAAAAACRAAKLODCTWLVNGDDLVVICESAGTQEDAAASLRFVTEAMTRYSAPODDPP	358
293	Db	CWLKVNAAAEQAGKNPRFLICGDDCTVIWKSAGADADKQAMVFASWMKVMGAPQDCVP	352
359	Qy	OPEYDLELITSCSSNVSAVHADSGKRVYLTDRPTVPLAPAAWETARHTVPVNSWLGNIIM	418
353	Db	QPKYSLEELITSCSSNVTSGITSGKPYFUTRDPRIPLGRCSAEGLYNESAAMIGVLIH	412
419	Qy	YAPTLWARILMTHFFSILLIAOQLEKALDCQIYGACYSEIPLDLPQIIRLRLHGLSAFSL	478
413	Db	HYPCLVSVSRVLAVHFMEQMLFEDKLPETVTFDWYKKNYTVPVDELPSIIAGVHGIEAFSV	472
479	Qy	HSVSPGEINRVASCIARKLGVPPLRVVHRHARSVRAKLLSOGGAAICGKYLFWNAVETKL	538
473	Db	VRVTNAEILRVVSQSLTDMTPPLURAWKKGARAVLASAKRGGHAHAKUARLEL-WHATSR	530
539	Qy	KUTP1PAASRLDLSGFWAGYS-----GGDIYHLSLR--ARPRHHHHH	580
531	Db	---PLP-----LDLKTSAVARYTTTFNYCDVYSPSGDVFTVPORLKOLEHHHHH	576

RESULT 11
US-09-921-397-108
; Sequence 108, Application US/09921397
; Patent No. US2002015148A1
; GENERAL INFORMATION:

```
; APPLICANT: HYBRIGENICS
; TITLE OF INVENTION: SID nucleic acids and polypeptides selected from a
; TITLE OF INVENTION: pathogenic strain of the hepatitis C virus and
; TITLE OF INVENTION: applications thereof
; FILE REFERENCE: B4809A - JAZ
; CURRENT APPLICATION NUMBER: US/09/921.397
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: EP 00402225.7
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 108
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-921-397-108

Query Match          30.4%; Score 623; DB 9; Length 127;
Best Local Similarity 89.8%; Pred. No. 2.1e-48;
Matches 114; Conservative 9; Mismatches 4; Indels 0; Gaps 0;

Qy 326 VICESAGTQEDASLRVFTTEAMTRYSPGPPQPEYDLELITSCSNVSVVAHDASGRV 385
Db 1 VICESAGVQEDASLRFTTEAMTRYSPGPPQPEYDLELITSCSNVSVVAHDGAGRV 60

Qy 386 YLITRDPTVPLARAWEETARHTPVNSLGNITMPTLWARMILMTHTFFSILLAOBLEK 445
Db 61 YLITRDPTVPLARAWEETARHTPVNSLGNITMPTLWARMILMTHTFFSILLARDOLEQ 120
Qy 446 ALDQIY 452
```

```
Db 121 AINCEIY 127
||:|:|

RESULT 13
US-09-921-397-106
; Sequence 106, Application US/09921397
; Patent No. US20020151484A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; TITLE OF INVENTION: SID nucleic acids and polypeptides selected from a
; TITLE OF INVENTION: pathogenic strain of the hepatitis C virus and
; TITLE OF INVENTION: applications thereof
; FILE REFERENCE: B4809A - JAZ
; CURRENT APPLICATION NUMBER: US/09/921.397
; CURRENT FILING DATE: 2001-08-02
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 106
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-921-397-106

Query Match          19.1%; Score 585; DB 9; Length 137;
Best Local Similarity 83.9%; Pred. No. 6.9e-45;
Matches 115; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

Qy 9 TWTGALITPCAABESKLPINALNSLLRHHNLVYSTTSRSASLRQKVTFFORLQVLDHY 68
Db 1 SWTGALITPCAABESKLPINALNSLLRHHNLVYSTTSRSACQKQKVTFFORLQVLDHY 60

Qy 69 RDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAQVRSLSRAVNHRSYVK 128
Db 61 QDVLKEVKAASKVKANLLSVEEACSLTPPHSAKSKFGYGAQVRSLSRAVNHRSYVK 120

Qy 129 DLLEDTDTPITTIMAK 145
Db 121 DLLEDSTVPTIDTIMAK 137

RESULT 14
US-09-294-121A-82
; Sequence 82, Application US/09294121A
; Patent No. US20020069422A1
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT; STUYVER, LIEVEN;
; APPLICANT: ROSSAU, RUDI; VAN HEUVERSWYN, HUGO
; TITLE OF INVENTION: PROCESS FOR TYPING OF HCV
; TITLE OF INVENTION: ISOLATES
; NUMBER OF SEQUENCES: 97
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIERMAN & MUSERLIAN
; STREET: 600 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/294.121A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/256,568
; FILING DATE: 18-JUL-1994
; APPLICATION NUMBER: PCT/EP93/03325
```

```
;
; FILING DATE: 26-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP/93/402,129.6
; FILING DATE: 31-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP/92/403,222.0
; FILING DATE: 27-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: CHARLES A. MUSERLIAN
; REGISTRATION NUMBER: 19,683
; REFERENCE/DOCKET NUMBER: 410.004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 661-8000
; TELEFAX: (212) 661-8002
; INFORMATION FOR SEQ ID NO: 82:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 113 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: be90
; POSITION IN GENOME:
; MAP POSITION: ns5
;
US-09-294-121A-82

Query Match 18.7%; Score 570; DB 9; Length 113;
Best Local Similarity 95.6%; Pred. No. 1.2e-43;
Matches 108; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 230 STVTENDIRVEESIQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNGCYRRCRASGVL 289
Db 1 STVTENDIRVEESIQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNGCYRRCRASGVL 60

Qy 290 TTSCGNTLTCYLKAAACRAAKLQDCTMLVNGDDLWVTCESAGTQEDAAASLRV 342
Db 61 TTSCGNTLTCYLKAAACRAAKLQDCTMLVNGDDLWVTCESAGTQEDAAASLRV 113

RESULT 15
US-09-899-082A-82
; Sequence 82, Application US/09899082A
; Patent No. US2002010638A1
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT; STUYVER, LIEVEN;
; ROSSAU, RUDI; VAN HEUVERSWYN, HUGO
; TITLE OF INVENTION: PROCESS FOR TYPING OF HCV
; ISOLATES
; NUMBER OF SEQUENCES: 97
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIERMAN & MUSERLIAN
; STREET: 600 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/899,082A
; FILING DATE: 06-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/378,900
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 08/256,568
; FILING DATE: 18-JUL-1994
; APPLICATION NUMBER: PCT/EP93/03325
; FILING DATE: 26-NOV-1993
```

```
;
; APPLICATION NUMBER: EP/93/402,129.6
; FILING DATE: 31-AUG-1993
; APPLICATION NUMBER: EP/92/403,222.0
; FILING DATE: 27-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: CHARLES A. MUSERLIAN
; REGISTRATION NUMBER: 19,683
; REFERENCE/DOCKET NUMBER: 410.004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 661-8000
; TELEFAX: (212) 661-8002
; INFORMATION FOR SEQ ID NO: 82:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 113 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: be90
; POSITION IN GENOME:
; MAP POSITION: ns5
; SEQUENCE DESCRIPTION: SEQ ID NO: 82:
;
US-09-899-082A-82

Query Match 18.7%; Score 570; DB 9; Length 113;
Best Local Similarity 95.6%; Pred. No. 1.2e-43;
Matches 108; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 230 STVTENDIRVEESIQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNGCYRRCRASGVL 289
Db 1 STVTENDIRVEESIQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNGCYRRCRASGVL 60

Qy 290 TTSCGNTLTCYLKAAACRAAKLQDCTMLVNGDDLWVTCESAGTQEDAAASLRV 342
Db 61 TTSCGNTLTCYLKAAACRAAKLQDCTMLVNGDDLWVTCESAGTQEDAAASLRV 113

Search completed: September 22, 2005, 14:58:37
Job time : 175 secs
```

This Page Blank (uspto)